CRANFIELD UNIVERSITY

STEPHEN JARVIS

A MISJUDGED APPROACH TO A HIGH ACCIDENT RATE: EXPLORATION OF ACCIDENT CAUSES AND INSTRUCTOR DECISIONS RELATING TO INEXPERIENCED GLIDER PILOTS

SCHOOL OF ENGINEERING

Ph.D. THESIS
Academic year: 2006 - 2009

Supervisor: Dr. Don Harris
February 2009

CRANFIELD UNIVERSITY

SCHOOL OF ENGINEERING

Ph.D. THESIS

Academic Year 2006-2009

STEPHEN JARVIS

A MISJUDGED APPROACH TO A HIGH ACCIDENT RATE: EXPLORATION OF ACCIDENT CAUSES AND INSTRUCTOR DECISIONS RELATING TO INEXPERIENCED GLIDER PILOTS

Supervisor: Dr. Don Harris

February 2009

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Abstract

Very little research has focussed on inexperienced pilots and there is a notion in literature and popular texts that such pilots are one of the safest groups. Despite this, 'early-solo' UK glider pilots were found to have a higher accident rate than any other group. Research was conducted in order to investigate the issues surrounding this. It was identified that accidents sustained by these pilots tended to be initiated by events in the approach and landing phases of fight, and caused by misjudgement of the approach path and landing flare. Most accidents to more experienced pilots were found to be different in all respects. It was subsequently found that instructors believed the highest accident likelihood to be associated with more experience pilots, in line with literature. It was also found that instructors wrongly believed that the 'approach' phase was the least likely in-flight phase to be associated with accident causes for low-hours pilots. Critical Incident Technique was used to investigate instructor decisions with regard to sending pilots solo. An initial model of the decision process was put forward. It was found that, with one critical exception, when events occurred on assessment flights that were similar to causal accident factors (from the accident analysis), instructors disallowed solo flight. An absence of potential accident factors was apparently insufficient to allow solo flight by itself; instructors required further evidence in order to confirm that students were ready to fly alone. Exceptionally, pilot performance in terms of the approach path did not appear to be a critical factor when instructors considered disallowing solo flight, highlighting a possible gap in the instructor decision process. It was recommended that further research be conducted to validate and extend the decision model, and that the approach phase be focussed upon more in both training and assessment.

Acknowledgements

Firstly huge thanks to my supervisor Don Harris, who became highly skilled at knowing exactly what I didn't want him to say... and saying it. Also for always telling me to write less, and less, and less, and for turning around my drafts in no time at all with accurate and worthwhile criticism despite being snowed under. Cheers Don!

Thanks is not enough for the legendry Paul 'The' Watson. Countless references to some unsung 'SME' or '2nd rater' give no clue to the amount of time Paul put into various parts of the project, in SME discussions, inter-rater tests, validation, etc. Even putting up with constant mundane discussion of fine details all the way to Buckminster GC and back on numerous occasions (probably shouldn't still have been discussing it at 900ft half way through a rolling circle though!). Thanks again Paul!

Big thanks too to other raters and SMEs who so kindly gave time to sit and work through tests of reliability and validity and various other tedious bits and bobs... Particularly to Delphine Tabary, Matt Ebbatson and Rebecca Wilson.

Thanks to all the other SMEs, members, instructors and managers at the various UK gliding clubs who made the work possible.

Lastly thanks to the British Gliding Association for providing the accident data.... even though the findings probably weren't what you wanted to hear!

CONTENTS

APPENDICES

List of Figures

List of Tables

List of Tables (continued)

List of Tables (continued)

Glossary of Notation and Abbreviations

AAIB	- Air Accidents Investigation Branch
AGL	- Above Ground Level
AOPA	- Aircraft Owners and Pilots Association
BASI	- Bureau of Air Safety Investigation
BGA	- British Gliding Association
CAA	- Civil Aviation Authority
CFI	- Chief Flying Instructor
CFIT	- Controlled Flight into Terrain
CIT	- Critical Incident Technique
CRM	- Crew Resource Management
df	- Degrees of Freedom
FAA	- Federal Aviation Administration
GA	- General Aviation
HF	- Human Factors
HFACS	- The Human Factors Analysis and Classification System
LOC-VMC	- Loss of Control in Visual Meteorological Conditions
NTSB	- National Transportation Safety Board
p	- Probability (statistical)
P1	- Pilot-in-Command
P2	- Second pilot or trainee
PA1	- Performance Assessment 1
PA2	- Performance Assessment 2
PIC	- Pilot in Command
PPL	- Private Pilots Licence
PPL (A)	- Private Pilots Licence - Aeroplanes
sd	- Standard Deviation
SE	- Standard Error
SME	- Subject Matter Expert
VFR	- Visual Flight Rules
VMC	- Visual Meteorological Conditions

Chapter 1 - Introduction

Gliding is a popular form of recreational and sport flying. In 2005 there were 9,000 active pilots in the UK and 20,000 licensed pilots in the USA (Jarvis & Harris 2007a). Evidence shows that accidents continue to occur annually; in a single five year period there were 117 non-fatal and 26 fatal gliding accidents in the USA alone (van Doorn & de Voogt 2007). It has been shown that 80% of these accidents were attributable to ‘pilot error’, including the majority of fatal accidents (van Doorn & de Voogt 2007). There is some suggestion in the statistics that certain flight phases contribute to more accidents than others, although this evidence is very limited. The ‘cruise’ phase has been shown to be associated with high numbers of accidents (van Doorn & Zijlstra 2006), although it is not clear what is meant by the ‘cruise’ phase for a glider, and the only other in-flight phases in this research were ‘landing’ and ‘take-off’. With the exception of the research already cited, there is no published research that looks at gliding accidents or safety.

Much research has focussed on other forms of recreational and light aviation, and particularly general aviation (GA). However it would be a mistake to simply generalise the results of these analyses across to gliding in order to inform safety programmes. Powered aircraft operate quite differently to gliders in every part of flight. Gliders launch in a very different manner to the way in which powered aircraft take off, they cannot sustain flight for long periods without the correct environmental conditions and they require different techniques and different controls in order to land. Accident statistics show that major causes of accidents in general aviation include ‘fuel management’, ‘poor selection of terrain for take-off’, and ‘go-around decisions’ (Wiegmann et al 2005). None of these are applicable to gliders. Glider pilots are arguably more vulnerable in other ways. They lack a number of general options when flying due to having no power available, for example they cannot ‘hold’ in order to wait for conditions to improve or for a runway to become available, and they cannot reject a poor circuit or approach by performing a ‘go-around’. It is therefore likely that there are accident themes in gliding that would not appear in powered aviation. For these reasons it is important that research takes place specifically for gliding, in order that remedial action can be taken to reduce accidents.

It has been found that only 5% of 'pilot-error' related gliding accidents are due to inexperience (van Doorn & de Voogt 2007). However the definition of inexperience is not given, and it is not explained how pilot inexperience was the 'cause' of these accidents. Popular literature from many forms of aviation asserts there is a band of moderate pilot experience which accounts for the highest accident rates (see Pratt 2000, Telfer 1993, Jenson 1995) and this band does not include highly inexperienced pilots. In gliding it has been proposed that early solo pilots are 'statistically' safer than other pilots (Piggott 1997).

Prior to any research being conducted, all UK accident data from 1997 to 2006 (British Gliding Association 2007b) were used to produce a number of simple frequency distributions of accident totals grouped by pilot experience. The experience level was expressed as the total number of hours gliding that the pilot had accumulated as pilot-in-command at the time of the accident. This is how pilot experience is expressed in the British Gliding Association accident database as well as many previous studies (see Aircraft Owners and Pilots Association - AOPA 2006, Booze 1977, Hasselquist 1999, O'Hare and Chalmers 1999, van Doorn & de Voogt 2007). Statistical analysis was not performed due to the fact that, in common with the flying totals, the accident database represented the entire population of accidents for the period, and therefore statistical tests would have been inappropriate (Ludwig 2005).

Two frequency distribution charts are shown. Figure 1.1 shows pilot experience grouped by brackets of 10 hours (up to 189.9 hours) and Figure 1.2 shows brackets of one hour (up to 9.9 hours). Each chart shows the accident counts for the ten year period 1997 – 2006 (inclusive) as well as the count for the most recent three year period (2004 -6).

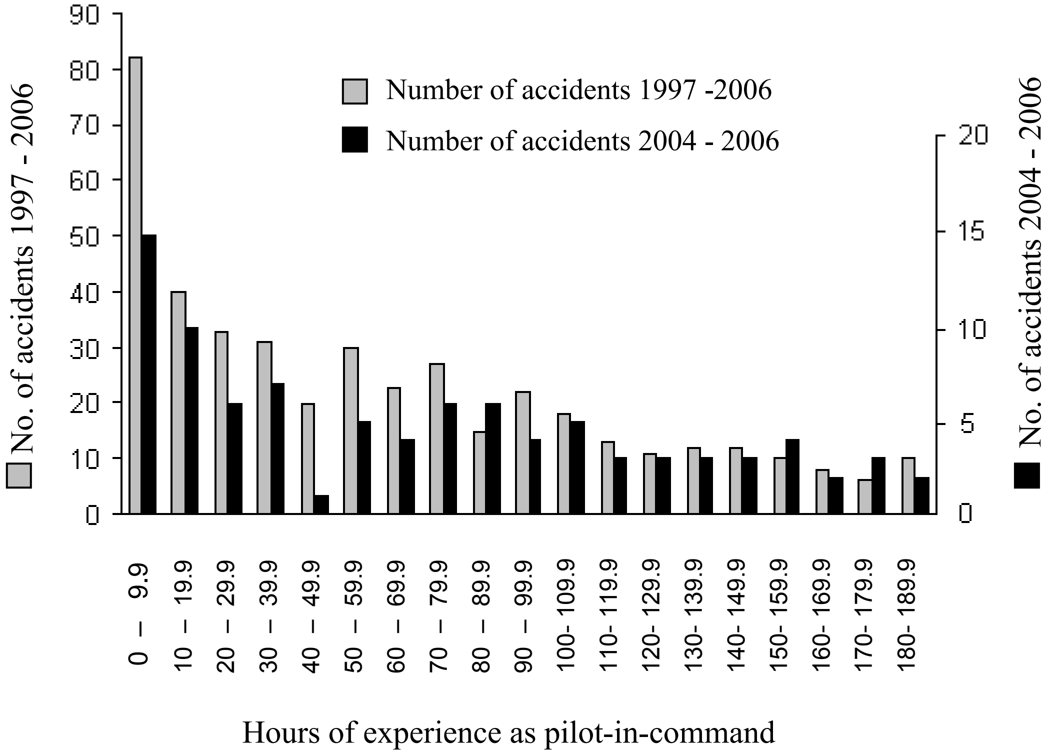


Figure 1.1. Frequency distribution of all UK gliding accidents grouped by pilot experience (10 hour brackets). Each block of ten hours experience is shown on the horizontal axis, and the number of accidents is on the vertical axis: 1997 – 2006 (left), 2004 – 2006 inclusive (right).

Figure 1.1 shows that the highest accident count belongs to the least experienced pilot group. Indeed when taken over a ten year period, the 'under-10' bracket has over twice the number of accidents of any other bracket. The most recent three year period (2004 – 2006) has the same general pattern as the ten-year period (1997 – 2006) and there is therefore reason to believe that the recent three-year period is representative of long-term accident trends in UK gliding.

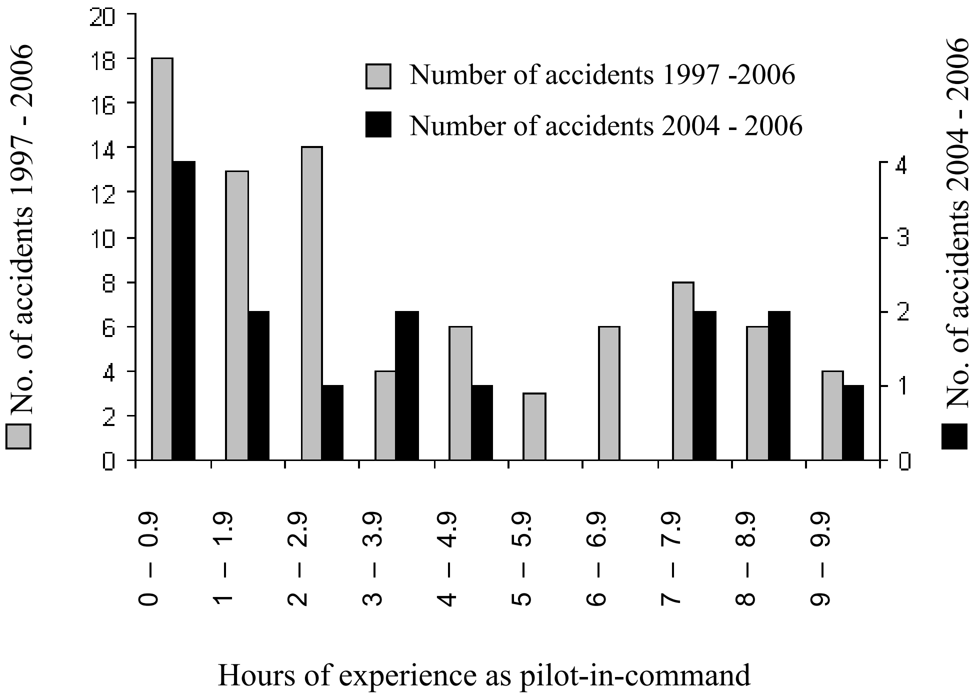


Figure 1.2. Frequency distribution of all UK gliding accidents grouped by pilot experience (1 hour brackets). Each block of one hour is shown on the horizontal axis, the number of accidents is on the vertical axis; 1997 – 2006 (left), 2004 – 2006 inclusive (right).

Figure 1.2 shows that proportionally more accidents occurred to pilots in their very first hour of flying. It can be seen that a recent three-year period (2004 – 2006) is broadly similar to the most recent ten-year period (1997 – 2006).

These simple charts suggest that there is a concerning trend. Raw UK gliding accident totals show consistently decreasing numbers of accidents with pilot experience, even when looked at in terms of single hours of flight experience, or when numbers are relatively small (2004 - 2006 only). This appears to be out of line with popular notions and indicates that, as a group, inexperienced UK glider pilots may require attention. Research is required into what factors cause these high accident totals in order that remedial action can be taken.

However to proceed with such research on the basis of this evidence alone would be premature. To claim on the basis of such simple data that low hours pilots are more likely to have accidents than experienced ones would involve the unsupported assumption that there is a constant exposure for all pilots (Li 1994). It cannot be known from accident counts alone whether this observed trend is simply a reflection of the amount of hours and launches flown by inexperienced pilots, compared to more experienced pilots. Inevitably the total number of flights taken across the glider population also decreases with experience. It is self-evident that more glider pilots get exposed to their tenth solo hour than their hundredth, since like any activity there will be an inevitable amount of participation attrition. Therefore the trend displayed in figures 1.1 and 1.2 could simply reflect this attrition rate, or even be opposed to it.

Instead of accepting the trend from the simplistic data analysis and proposing research aimed at investigating it, there is first a need to review literature around this area and to obtain and analyse appropriate data to confirm or discount the notion that inexperienced glider pilots are a particularly vulnerable group.

Chapter 2 - Analysis of accident totals against exposure measures for low hours glider pilots

This study was published as:

Jarvis, S. Harris, D. (2007b), *Accident Rates for Novice glider pilots Vs pilots with experience.* Aviation, Space and Environmental Medicine Vol 78: p.1155-1158

Introduction

A best selling UK gliding textbook claims that "Statistically, your first solo flights are safer than most other flights because the pilot has recently revised all the critical situations and is in good flying practice" (Piggott 1997). No reference is given for this statistic.

The quote corresponds to the popular assertion referred to in the previous chapter that pilots with a moderate amount of flying experience, rather than the least experienced pilots, have the highest accident rates, whereas low-hours pilots have a low risk of being involved in an accident. In powered aircraft flying, where most research has been conducted, there are many claims to this effect. Jenson (1995) relates that there is a "universally well-known high accident rate for pilots with between 100 and 300 total flying hours", while Olsen and Rasmussen (1989) state that the greatest risk to pilots occurs at "around the 100 hour mark". However the source of these figures is not given. Telfer (1993) claims that low-hours pilots are less represented than more experienced pilots in accident statistics, and are therefore a relatively safe group.

No studies of glider statistics are available to support claims regarding the relative safety of inexperienced glider pilots, but other forms of recreational aviation, particularly general aviation, have been studied.

In U.S. general aviation it has been found that cumulative flight time experience increases the rate of accident involvement, peaking between 1001 and 2000 hours, when the rate begins to reduce (Booze 1977). Pilots with 10 hours or less have been shown to have the lowest accident rate of any experience group; over ten times less than the 1001 – 2000 hour group (Booze 1977). This particular study has been cited by many authors (e.g. O'Hare et al 2001, Telfer 1993).

Many general aviation textbooks report accident figures in a way that supports the trend of these claims. Wells (1992) states that the highest proportion of accidents take place to pilots with between 100 and 499 hours. Pratt (2000), using a simple count of accidents, observed that out of 300 UK fatal accidents: "Less than 5% of the fatalities were student pilots; far more instructors and professional pilots were involved in fatal accidents" (Pratt 2000).

The notion that highly inexperienced pilots are relatively safe has attracted theoretical explanations based on psychological mechanisms such as increased risk-taking behaviour, overconfidence and an increased perception of 'invulnerability'. Booze claims that "Overconfidence and lack of vigilance by high time pilots have been cited as possible contributors to the situation" (Booze 1977) although no references are given. Some more recent studies support the existence of such mechanisms.

Telfer (1989) describes the findings of such a study, often referred to as 'The Australian Study' (Telfer & Ashman 1986);

> At the beginning of their training, student pilots recognised that their judgement could be a source of danger to themselves and others. After accumulating flying hours over half a year they were less prepared to concede that their judgement could be defective. Such a perceptual shift might correspond to the high accident rate of pilots who have total flying hours in the range of 100 – 300 hours. They may have a similarly optimistic view of their invulnerability (Telfer 1989).

In a simulated decision making test of mixed pilots, O'Hare (1990) found that experienced pilots were more likely to choose to accept a marginal VFR flight than less experienced ones, although the results only tended towards significance. The data suggested that greater experience led to a higher risk utility and a greater sense of ‘invulnerability’ (O’Hare 1990). Wichman and Ball (1983) found evidence that “aviators with more experience and exposure develop stronger self-serving biases” (from O’Hare 1990).

There is a growing body of evidence that appears to refute these claims, and suggest relatively high accident rates exist among inexperienced pilots. In a study of homebuilt aircraft accidents for the single year of 1993, Hasselquist and Baker (1999) found that student pilots were involved in 18% of the 52 fatal accidents recorded, although no exposure measure was collected. Considering that the mean total flight time of homebuilt aircraft pilots (from the accident data) was 3088 hours, it would be very surprising if student pilots accounted for 18% of the total ‘non-accident’ flying hours. Hence this suggests a high accident rate among student pilots. A high incidence of fatal accidents among ultralight pilots with less than 40 hours experience on the specific aircraft type (make/model) has been found (Pagan, De Voogt & Van Doorn 2007). Although this effect disappeared when overall experience was factored in, there was no hint that relatively more experienced pilots were at greater risk than low-hours pilots. A recent study by The Flight Safety Foundation pointed out that “Pilots with fewer than 200 hours total time are the most vulnerable and those with fewer than 10 hours in make and model are more vulnerable still” (AOPA 2006). These studies hint at the possibility that low-hours pilots may in fact have a higher relative accident rate, or at least have a rate no lower than the average for the pilot population.

Unfortunately, all of these studies are based simply on counts of accident data. Without properly accounting for the amount of flying that pilots of different experience groups are exposed to, accident studies cannot fully support or revoke claims made regarding accident involvement and experience. This can only be done by properly factoring in an appropriate measure of the amount of flying that each group undertakes (an exposure measure). A failure to do so leads to the unsupportable assumption that the amount of flying is constant for different groups (Li 1994).

The difficulty of gathering appropriate exposure data to revoke or support such claims remains a problem with the research literature. Accident data are relatively easy to get access to and often contain considerable detail including pilot experience information. However accident counts alone cannot show the likelihood of accidents occurring. Neither Wells (1992) nor Pratt (2000) gave any indication of overall numbers of pilots in any categories, nor the amount of flying they had done. Therefore even if their figures are correct, they may represent an inaccurate picture of the situation. For example in the year 2000 a UK student pilot only needed to log a minimum of nine hours solo flying for the qualification of a PPL (A), meaning student pilots (as pilot in command) had very few solo hours of exposure. This total may be very low compared to other pilots, and it is possible that Pratt's claim that fewer than 5% of fatalities were student pilots could in fact equate to a very high accident rate in comparison to other pilot experience groups.

Hence these sorts of studies cannot provide objective comparisons of accident rates between particular groups (e.g. inexperienced against experienced pilots). This can only be done by factoring in a measure of the amount of flying that each group is exposed to (an exposure measure), meaning a figure representing the amount of flying done by each pilot group, recorded either as numbers of hours flown or number of flights. Such data is not easily available. In recreational aviation the only records of hours flown by pilots in particular experience groups are contained within an individual pilot's logbook (O'Hare & Chalmers 1999). This makes it very difficult to produce appropriate measures of exposure, and helps to explain why few studies use them.

It is important to select an appropriate measure of exposure. According to Li (1994) studies often incorrectly use 'crashes per pilot per year' as an exposure measure, which would fail to give accurate information about flight time. An effective exposure measure should include a factor equating to the amount of accident-free flying done by groups of interest, in order to calculate the relative frequency of accident occurrence for true comparisons between the groups. Booze (1977) factored in the numbers of pilots in each group of interest, as obtained from medical records, and rates were constructed by comparing the number of 'accident' pilots in each group by the number of 'non-accident' pilots. Li (1994) has effectively demonstrated that the effect shown by Booze

(1977) was reversed when controlled for by pilot-flight-hours per group, rather than simply the theoretical number of active pilots.

This is supported by other studies using measures of exposure that have shown that inexperienced pilots are more vulnerable to pilot error accidents. Borowsky (1981) used comprehensive records kept by the US Navy in order to collect accurate figures for pilot exposure and accident involvement. It was found that for US Navy pilots there was a significant decrease in accidents involving ‘pilot error’ (as defined within the database) as total flight hours increased. The only recent study of recreational aviation to use a sophisticated measure of flight time exposure was carried out by O’Hare and Chalmers (1999), who produced the exposure measures from a large study of pilot logbooks in New Zealand. It was found that when controlled by measures of flying hours per year, the proportion of accidents in the 100 to 200 and the 100 to 300 hour brackets were in fact slightly lower than the general trend across all accidents.

No study investigating the accident rates among glider pilots has been undertaken. The only study so far to look at gliding found that of 114 accidents judged to have been caused by ‘pilot error’, only six were attributed to inexperience (van Doorn & De Voogt 2007). However no comparable accident rates were produced for the inexperienced group. Apart from this study, no recent accident statistics exist for gliders (van Doorn & De Voogt 2007).

Research aims and objectives

Most research in recreational aviation has been shown to be limited in terms of its ability to objectively inform of accident rates between pilot groups, and no study that makes appropriate use of exposure data has focussed on very early solo pilots. The main reason for this may be the difficulties involved in producing exposure measures in this field.

However the small amount of valid research of aviation accident databases does not support the popular assumptions about the existence of a zone of moderate pilot experience that is particularly high risk, nor is there any good evidence to support notions that early solo pilots are relatively safe. On the contrary, the studies tend to point towards the possibility that inexperienced pilots are relatively more vulnerable than others, which would mean that the popular notion could be dangerously misleading. Despite this, no accident studies have specifically looked at investigating accident risks among very early solo pilots in recreational aviation or gliding, by factoring in measures of flight exposure.

More such work does need to be done in order to properly inform pilots and instructors of the risks involved in early hours flying and establish the effectiveness of training and instructor judgement of student pilot suitability for solo flying.

Claims such as that by Piggott (1997) cannot be justified from the literature, and the small amount of valid research points in the opposite direction. This study aims to produce a valid estimate of the accident rate for very inexperienced UK glider pilots controlled for flight exposure, along with comparable data for the general pilot population.

Method

Main data collection

The amount of annual UK gliding was collected from 'annual statistics' published by the British Gliding Association (British Gliding Association 2004, 2005, 2006a). A sample of daily UK gliding log sheets (flying logs) were then collected from gliding clubs in order to estimate the proportion of annual flying attributable to very inexperienced pilots. In gliding, as in general aviation, the only individual pilot records are kept in pilot logbooks. However due to the particular operation required to launch gliders, many UK clubs keep daily log sheets of all launches, and most keep a log of hours flown by each glider, including the name of the pilot. Such log sheets vary in quality between clubs; but all give a list of daily launches and most contain individual flight times. These could therefore be used in order to sample the amounts of accident-free flying done, i.e. a measure of exposure.

Three UK gliding clubs took part, all of which had active training programs, used both aero-tow and winch launching methods, and kept daily records with an adequate amount of detail from the previous three years.

Accident counts were taken from the UK database of all gliding accidents and incidents from 1997 to 2006 (British Gliding Association 2007b). This contained details including age and experience (in hours) of the pilot in command, aircraft type, severity of injuries, damage to the glider and a description of what happened. All accidents where damage or injury were categorised as minor or above were included in the analysis. Incidents were not included. Additionally, ground-handling accidents (such as towing out winch cables, or pulling gliders out of hangers) were not included. These types of accidents made up 15.1% of the total in the database from 2004 to 2006. A small number of glider accidents did not include any data about pilot experience. This group made up 6.3% of the accidents from 2004 - 2006. There was no reason to suggest this small group of accidents was biased towards any particular experience group. Some were clearly experienced pilots due to the nature of the flight description, e.g. a competition or long cross-country flight, whereas others were less experienced pilots, e.g. containing descriptive comments such as "with the aim of completing a five hour badge flight" or "second flight on type". Since there was no reason to assume a systematic variation in terms of pilot experience, this small group of accidents was dropped from the analysis. Two accidents were subject to AAIB investigation and no information was available at the time of analysis, hence these were also dropped.

Flying hours were used as the measure of pilot experience, in line with many previous studies (AOPA 2006, Booze 1977, Hasselquist 1999, O'Hare & Chalmers 1999, van Doorn & de Voogt 2007). Two experience brackets were decided upon for this investigation:

1. Solo pilots with only one hour or less of experience as pilot-in-command (PIC). Therefore this only included those taking their first few solo flights. This will be referred to as Group G-1.
2. Solo pilots with 10 hours or fewer PIC experience (includes those in group 1). This will be referred to as Group G-10.

These groups both represent very inexperienced pilots who would have recently been involved in training and instruction. There were a number of firm reasons behind these specific groups being selected.

Firstly, the frequency distributions of accident totals for the ten year period 1997 - 2006 showed that twice as many accidents occurred to pilots with ten hours or fewer experience (G-10) as any other ten hour experience bracket (Chapter One, Figure 1.1). The pattern for 2004 - 2006 showed a similar general trend, with the 0 – 10 hour bracket having the highest accident total. The finer distribution within the ten hour bracket showed more accidents in the very first hour than any subsequent hour (Chapter 1, Figure 1.2). As previously discussed, these figures were determined from accident totals only, and so could not be used to draw conclusions as to the likelihood or proportions of these accidents occurring, since the accident distribution may simply be a reflection of the distribution of the amount of flying done in each bracket. However, the distributions gave a clear justification to collect exposure data to clarify the meaning of these findings for groups G-1 and G-10.

Secondly, the first major UK gliding qualification (the BGA Bronze Badge) includes a requirement of ten hours of flying time or 50 launches, and begins the process of restrictions being lifted on pilots flying out of gliding range of their base airfield. It was therefore determined that ten hours was a significant experience level in UK gliding. However it is possible that despite having fewer than ten hours, such pilots may have accrued many flight cycles. Although not used in this research due to lack of complete records, flight cycles could be argued to be an important measure of experience in gliding where many flights are short, and the majority of accidents occur on landing (van Doorn & de Voogt 2007). This supports the need to investigate a group such as the G-1 group which would be certain to have a low number of flight cycles.

Thirdly, gliding clubs do not record or update information about individual pilot experience (apart for instructors, for regulatory reasons). Therefore although daily gliding log sheets do identify pilots, they do not indicate the experience of those pilots. To do this, pilot information would have to be constantly updated and in some cases sought from years for which no records existed, or sought from other clubs. Many pilots fly at more than one club, but not normally as pilot-in-command unless they have some solo experience, normally at least a BGA 'Bronze' badge (meaning they had flown 10 hours solo, or had at least 50 solo flights). For these reasons pilot experience could only be accurately assessed from log sheets if the pilot in command had less than about ten hours solo. Such pilots could be identified on the log sheets (usually by club members or office staff) and their total flying experience traced back through archive log sheets where necessary. Such assessments would be increasingly unreliable with pilots of greater experience levels.

Exposure measure

The proportion of annual flying exposure was estimated from a sample of daily gliding log sheets. Three lists of 45 random annual dates were generated, labelled List A, List B and List C and checked for a reasonably representative monthly distribution. In order to guard against systematic sampling error, a counter-balanced sampling strategy was used. For a given club, list A was applied to the year 2004, B to 2005 and C to 2006. At the next club this order was rotated one place, meaning that List C was applied to 2004, A to 2005 and B to 2006. This rotation continued for the third club. No cluster or stratified sampling was used from this point, because although the summer season sees the greatest number of hours and launches, there is no reliable information about the distribution of various types of flying that occur across the year. For example it is likely that the distribution of serious cross-country flying and racing is heavily clustered around July and August, whereas the distribution of training and local flying could be flatter since it does not require thermal conditions. On the other hand, glider aerobatic flying often decreases in the summer months due to the thermal activity and increased local glider traffic. Since these distributions are not known, the best sampling strategy was a purely random one, with enough cases to adequately represent each month of the year.

A high degree of confidence was required for the sample, and so a calculation was made to ensure that its mean was highly likely to be within 0.1 of a standard deviation from the population mean (to a 95% level of confidence). A sample of at least 385 days of data was required to provide an estimate with a 95% level of confidence that the measures for the flight exposure data would be within 0.1 standard deviations of the population mean, prior to calculating further confidence intervals (Hays 1994). In order to achieve this, 45 days from each of the three years were sampled from each club, making a sample total of 405 days (20 days above the minimum calculated). The total number of hours and launches flown by the G-10 and G-1 groups was summed for each daily log sheet, along with the total number of hours.

Further triangulation of exposure measure:

The only UK gliding qualification with a requirement to attain a specified number of gliding hours is the BGA Bronze Badge, requiring 10 hours of solo flying experience and twenty flights, or just fifty flights (British Gliding Association 2001). Figures were published by the BGA from 2006 onwards giving the number of Bronze badges attained in the UK during 2005 and 2006 (British Gliding Association 2006a, 2007a). These figures were used to triangulate the exposure estimate by multiplying the number of Bronze badge qualifications per year by 10 and comparing to the annual hours of the 'G-10' group estimate.

A further requirement of the Bronze badge is two solo flights of at least 30 minutes duration. Therefore it would be possible to complete 50 flights with only five hours experience (2 x 30mins and 48 x 5mins). Equally it would be possible to achieve 10 hours flying without amassing 20 flights. However this would require a very inexperienced pilot to exploit good soaring conditions, since even from a high 3000ft aero-tow a glider with a normal sink rate of 200ft/min would be on the ground in 15 minutes without the use of lift. Therefore whereas not all pilots qualifying for a Bronze badge will have exactly 10 hours, the variation either side will be relatively small for the majority of claimants. The expectation therefore is that if the G-10 group exposure calculation were accurate (in terms of the number of hours flown), it should not be too far from the number of annual Bronze badges multiplied by 10 (for hours). Naturally some pilots drop out of the sport without reaching the Bronze badge, and therefore the direction of any expected difference should show as a slightly higher number of hours in the exposure measure than this bronze badge calculation.

Treatment of data

The mean number of launches and hours flown per day (for 2004 - 2006) was calculated for groups G-1 and G-10 from the exposure data collected. The mean number of hours and launches for all flying done (also as a daily figure) from 2004 to 2006 at the three clubs involved was taken from the BGA annual statistics (British Gliding Association 2004, 2005, 2006a). This meant that an accurate figure could be known for this measure, rather than a sample estimate. Having estimated figures of flying hours and launches for the low-hours pilot groups and derived the same figures for all pilots, the flight exposure measure for G-1 and G-10 pilots could be produced for the three year period. This was required in order to enable comparison with accident totals derived from the same three year period. Three estimates of flying exposure for inexperienced pilots (related to the period 2004 - 2006) were produced; an upper and lower estimate based on the 95% level of confidence, and a mid-point between the two (a mean estimate).

Accident rates were derived by dividing the total number of accidents for each group (from the BGA accident database: British Gliding Association 2007b) by the estimated number of hours and launches flown by that group over the same period. The overall accident rate (for all pilots) was obtained from known accident numbers and known totals for the amount of flying undertaken, from the data provided by the BGA.

Results

The results for this study are in two sections. Firstly the statistics for the exposure measures are displayed and secondly the accident rates are shown.

Section 1 - Exposure measure

The mean number of launches per day per club from BGA population parameters was 29.5 launches. Because these were necessarily calculated from annual totals, rather than daily totals, standard deviations could not be produced. Although not used, the same figures estimated from the sample data of 405 days flying were 28.7 launches/day (sd = 32.5) and 13.6 h/day (sd = 22.3). This suggested that the sample data were a good representation of daily UK gliding movements for that three-year period.

Table 2.1 shows the number of launches and hours collected from each group from the sample (405 days) along with the mean daily total and the standard deviation.

Table 2.1. Statistics for the 405 days of data collection.

	Total	Mean per day	**Std Dev**	**Std Error**	**Upper bound**	**Lower bound**
G-1 Launches (n)	54	0.13	0.5	0.025	0.18	0.08
G-1 Hours (h)	8.62	0.02	0.09	0.004	0.03	0.01
G-10 Launches (n)	209	0.52	1.2	0.06	0.64	0.4
G-10 Hours (h)	84.53	0.21	0.62	0.031	0.27	0.15

Using the mean totals for all pilots (calculated from BGA parameters) along with the statistics (Table 2.1) the proportion of exposure was calculated for the two groups, in terms of launches and hours, as a percentage of UK gliding activity. These proportions were used to estimate the amount of flying done across the UK for the period 2004 to 2006 inclusive (table 2.2).

Table 2.2. Estimated number of hours and launches taken by low-hours glider pilots in both groups and the associated percentages of flying done in the UK (2004-2006) with 95% confidence intervals.

		Upper 95% CI	Mean	Lower 95% CI
Group G-1	**Estimated number of launches taken**	5852	4276	2700
	Estimated total hours flown	896.16	633.52	370.88
Group G-10	**Estimated number of launches taken**	20372	16550	12727
	Estimated total hours flown	8050.23	6212.14	4374.04
Group G-1	**Percentage of launches taken (%)**	0.6	0.5	0.3
	Percentage of hours flown (%)	0.2	0.15	0.09
Group G-10	**Percentage of launches taken (%)**	2.2	1.8	1.4
	Percentage of hours flown (%)	1.9	1.5	1

Triangulation of the exposure measure estimate.

The number of UK Bronze badges attained in 2005 was 200 and in 2006 it was 196 (British Gliding Association 2006a, 2007a). Based on these data, it was expected that the number of hours flown by the Group G-10 was close to 1,980 per year (5940 for the three year period), and probably higher to account for the pilots dropping out prior to completing the Bronze badge. The exposure figure obtained from sample data (table 2.2) for the three year period 2004-2006 was 6212.4 hours (2071 hours per year). Thus the exposure estimate and the estimate from the Bronze badge figures are very comparable, verifying the exposure measure.

In total there were 268 glider accidents causing damage or injury (2004 – 2006) of which 16 were accidents that involved solo pilots with 10 solo hrs or fewer and six involved solo pilots with one hour or less. Table 2.3 shows the estimated numbers of accidents per launch (and launches per accident) for the two inexperienced pilot groups, as well as the overall accident rate for glider pilots. Table 2.4 shows the same information, but for hours flown rather than launch numbers.

Table 2.3. Estimated UK glider accident rates per launch for 2004 – 2006 inclusive. Bracketed figures give the number of launches flown per accident. The overall accident rate per launch is included for comparison, calculated from BGA annual flight safety statistics. Groups G-1 = pilots with one hour or less experience, G-10 = pilots with ten hours or less

	Accident rates	Upper 95% confidence interval	Mid-Point	Lower 95% confidence interval
Group G-1	**Accidents/launch**	**0.001**	**0.0014**	**0.0022**
	(Launches per accident)	(976)	(713)	(450)
Group G-10	**Accidents/launch**	**0.0008**	**0.001**	**0.00125**
	(Launches per accident)	(1,274)	(1,034)	(785)
UK Overall	**Accidents/launch**		**0.0003**	
	(Launches per accident)		(3,534)	

Table 2.4. Estimated UK glider accident rates per hour flown for 2004 – 2006 inclusive. Bracketed figures give the number of hours flown per accident. The Overall accident rate per hour is included for comparison, calculated from BGA annual flight safety statistics.

	Accident rates	Upper 95% confidence interval	Mid-Point	Lower 95% confidence interval
Group G-1	**Accidents/hour (h)**	**0.0067**	**0.0095**	**0.016**
	(Hours per accident)	(149)	(106)	(62)
Group G-10	**Accidents/hour (h)**	**0.002**	**0.0026**	**0.0037**
	(Hours per accident)	(503)	(388)	(273)
UK Overall	**Accidents/hour (h)**		**0.0006**	
	(Hours per accident)		(1,590)	

It can be seen that the accident rates for both groups are higher than the overall accident rates, whether looked at in terms of flight time or flight cycles.

Discussion

Simple frequency distributions based on raw accident counts showed consistently decreasing accidents with increasing experience (see Chapter 1, figures 1.1 and 1.2). However without knowing whether this trend was simply a reflection of the numbers of flights or hours flown by pilots of those experience levels, conclusions could not be drawn regarding the rates of accidents. This study used two measures of flight exposure to estimate the true accident rates of pilots with one hour or less, and 10 hours or fewer.

Taking the lowest estimate among inexperienced pilots (using the upper 95% confidence interval) it can be seen that these pilots sustained much higher accident rates than the overall UK rate for all glider pilots (tables 2.3 and 2.4). By this lowest estimate, the G-1 group were involved in over three and a half times the number of accidents per launch than average and over ten times the number of accidents per hour flown. For the G-10 group the figures were slightly better at just under three times the number of accidents per launch than the overall UK rate for all glider pilots, and just over three times the number per hour flown. Since this is the lowest estimate, it is probable that the real situation is even worse. Using the most likely estimate (the mid-point between the upper and lower confidence intervals, tables 2.3 and 2.4) gives 4.96 times the number of accidents per launch and 15.1 times the number per hour for the G-1 group, and 3.42 times the number of accidents per launch and 4.1 times the number per hour for the G-10 group.

Despite being in a different area of recreational aviation, the results support the observation by AOPA that pilots with fewer than ten hours are vulnerable (AOPA 2006). Findings of O'Hare and Chalmers (1999) are also partially supported, in that after the early stages of solo flying the accident rate falls as experience is acquired. However since the current study only used exposure data for pilots with a few hours, it cannot provide information with regards to the 100 - 300 hour experience bracket.

The high hourly accident rate for the G-1 group could be partly accounted for by those pilots flying relatively short flights. It is likely that exposure to flight cycles presents a greater risk to pilots than exposure to hours flown, because of the increased risk presented by the takeoff and landing phases of each flight cycle. Analysis of 143 glider accidents in the USA observed that 59.5% of the accidents occurred in the landing phase (van Doorn & de Voogt 2007).

The reason for early-hours pilots doing short flights is that during their early solo flying, pilots are restricted to flying circuits in benign, often unsoarable conditions. Against this background the high accident rate should be of great concern because inexperienced pilots are the most closely supervised and restricted of all glider pilots. This period involves being checked for solo competence by an instructor before flight, particular for very early solo pilots. In addition restrictions are put on the pilots' activities. These would generally include only being allowed to fly within gliding range of the airfield, not flying if there is a marked change in conditions, only using a specified launch method, and being restricted to flying basic glider types. Furthermore early solo pilots are only allowed to fly in conditions that are deemed safe, including weather, times of day, direction of the sun on landing, visibility, traffic density and landing directions (runs). Given all of these precautions and the fact that these pilots fly the most basic flights in the easiest conditions, it is of concern that their accident rate remains so high. This may indicate that some pilots are being allowed to fly solo before they are ready.

Whereas powered aircraft pilots have the option of escaping from, and postponing many situations by applying power, glider pilots are forced to confront situations directly, even when beyond their experience. For example they do not have the option to 'go round' from a poor approach, divert due to poor weather, hold for other traffic or climb to avoid obstacles. Therefore glider pilots require the ability to quickly resolve these problems, without the possibility of making more time available. It may be that many pilot weaknesses are not exposed during the controlled training environment, possibly because some situations cannot be easily engineered by instructors. It is therefore possible that these abilities are not practiced or tested enough before pilots are sent solo.

Conclusion

No previous study has uncovered such a strong effect on low-hours pilots. This may be due to the type of study; direct focus on low-hours pilots has not previously been done using a suitable measure of exposure. It may also be due to the difference in the operation of gliders, as opposed to other forms of aviation.

The findings of this study should be used to inform the gliding community that contrary to some popular assumptions, early solo pilots have a relatively high risk of being involved in an accident. Further research is now required in order to ascertain what the specific problems are that cause low-hours pilots to have such a disproportionate accident rate. Until more is known about the specific types and causes of accidents occurring to these pilots and whether they are different to the overall accident trend, remedial action cannot be properly focussed.

This study suggests that important safety-critical learning takes place after pilots are sent solo (progressively decreasing the risk of an accident). It is not possible to know whether such learning could be achieved while still under instruction and so there is a question of whether some pilots are being sent solo too early.

The findings from this study contradict claims made about the relative safety of low-hours glider pilots. Flight exposure estimates show that inexperienced UK glider pilots are more likely to have an accident than other more experienced pilots whether the figure is looked at by flight cycles (per launch) or by hours flown. The findings show that claims such as “statistically, your first solo flights are safer than most other flights” (Piggott 1997) are wrong.

Chapter 3 - UK Gliding Qualifications

It has been robustly established that there is a safety issue surrounding highly-inexperienced UK glider pilots. Before looking further into this issue, it is necessary to outline the basic UK gliding qualifications applicable to the subsequent studies.

Gliding in the UK

The governing body for gliding in the UK is the British Gliding Association (BGA). The BGA not only provides leadership and representation in UK gliding but also acts as a self-regulator, supplemented where appropriate with delegations from the Civil Aviation Authority. Therefore the BGA has a large role to play in the safety of gliding in the UK, and to this end has committees dedicated to safety and instruction. BGA responsibilities include the setting of standards and progress levels for pilots, instructor training and standards, and the investigation and documentation of accidents.

Solo qualification

UK glider pilot trainees follow a training syllabus laid down by the British Gliding Association (BGA). There is no set minimum number of flights or hours required before a pilot can go solo, but there is a minimum of 20 launches that is required before a 'test' can be taken to obtain a certificate or badge. The accrediting document issued to British glider pilots is the Gliding Certificate, which is issued on completion of their first solo. This certificate is endorsed with the A and B badges shortly after the first solo. The tests for these are as follows (source: 'The BGA Laws and Rules' (British Gliding Association 2008a).

- ‘A’ BADGE: One solo circuit in a glider or motor glider in unpowered flight after the launch, followed by a satisfactory landing. An appropriate level of knowledge of rules of the air and local airspace restrictions must be demonstrated to the supervising instructor at the time of the first solo flight

- ‘B’ BADGE: A soaring flight of at least five minutes, at or above the previous lowest point after launch, followed by a satisfactory landing.

Post solo qualifications

Once a glider pilot has gone solo and achieved the A and B badges, he or she can work towards their Bronze ‘C’ badge (commonly known simply as ‘The Bronze Badge’) which is the first major gliding qualification in the UK. In order to obtain this the pilot must achieve two observed soaring flights, 50 total flights (or 10 hours and 20 flights), pass a flying test with a senior instructor and successfully take an exam covering subjects such as air law, principles of flight and navigation. The Bronze Badge is therefore similar to the Private Pilot License, in that it is the major qualification obtained following solo experience, tests and exams. Although a UK gliding license does exist, it is voluntary and not commonly applied for apart from by pilots looking to fly abroad, indeed this was one of the main reasons for its implementation (British Gliding Association 2006b). The license requires the holder to have a Bronze Badge and an endorsement that they are capable of cross-country flying. The syllabus for trainee glider pilots and for the Bronze badge is not as thoroughly defined as that of the Private Pilots License (PPL). Unlike the PPL, there is no minimum number of dual flying hours required and although many exercises are specified in the glider pilot trainee syllabus, no evidence needs to be provided that all these have been performed in order to obtain a Bronze Badge.

Chapter 4 - Fundamental Accident Analysis

The research from chapters four and five was published as:

Jarvis, S. Harris, D. (2008), *Investigation into Accident Initiation Events by Flight Phase, for Highly Inexperienced Glider Pilots*. International Journal of Applied Aviation Studies Vol 8, Number 2. FAA Academy.

Introduction

Jarvis and Harris (2007b) used flight exposure estimates based on sample data from a three year period along with BGA statistics and simple accident counts, to show that very inexperienced solo glider pilots (defined as those with fewer than 10 hours experience as pilot in command) had twice the number of accidents per launch and three times the number per hour flown than their more experienced counterparts (see chapter 2). Important questions were generated about what lies behind this. Pilot inexperience alone is not adequate as an explanation for such a phenomenon. With such a small amount of research in the field of gliding, deeper explanations are required about how and why inexperienced pilots sustain these high accident rates if action is to be taken to address them.

The aim of the present study was to investigate fundamental accident causation in UK gliding, in order to establish the proportions of accidents that were pilot related. Particular focus was placed upon inexperienced pilots (those with up to 10 solo hours). This was done for two reasons. Firstly to test the hypothesis that pilot related accident causes were driving the high accident rate, and secondly to lay the groundwork for subsequent studies to look deeper into the exact nature of the reasons behind the earlier findings (chapter 2). Accidents identified as ‘pilot related’ were used subsequently as the primary source of data.

There is wide acceptance that human factors are the most significant source of failure in aviation as well as in other safety critical industries (O’Hare 2000) and the estimated proportion of accidents caused by pilot factors remains high for all types of aviation. The figures reported from aviation related research over the last quarter of a century have remained consistent. Feggetter (1982) reported that the figure was approximately 70%. More recent studies have reported figures of the same magnitude, for example 71% (O'Hare et al 1994); 70% (The Bureau of Air Safety Investigation – BASI 1996); and 78.6% (Aircraft Owners and Pilots Association - AOPA 2006). Similar findings were reported after examining 143 glider accidents in the US over a single five year period. The majority of accident causes (79.7%) were designated as pilot related, 14% were designated as weather related, 4.9% malfunction (mechanical) and 1.4% were undetermined (van Doorn & de Voogt 2007). No research has split these figures by pilot experience level, nor analysed them specifically for low-hours pilots.

On initial consideration, there would appear to be no reason why the proportion of mechanical and external factors would change among pilots of different experience levels. However, this should not be assumed. The aims of this study were therefore:

- To identify all UK gliding accidents from 2002 - 2006 that were 'pilot related' in order for these to be further analysed in later studies.
- To investigate the proportion of UK gliding accidents that were pilot related, and then to investigate whether the inexperienced group of pilots, defined as those with 10 or fewer hours as pilot-in-command (PIC) had a higher proportion of such accidents.
- To investigate the categories of damage and injury, in order to assess whether or not inexperienced pilots suffered a different level of severity in accidents to more experienced pilots.
- To investigate whether there is evidence that weather conditions (specifically wind strength) could have been a factor in the high proportion of accidents by low-hours pilots.
- To identify the specific set of 'pilot-related' accidents occurring to student pilots under instruction. This was done in order that later comparisons could be made between types of accidents occurring to students before and after they had gone solo.

Method

Guidelines were constructed by a group of subject matter experts based on the work of Hollnagel (1998). These were used to identify all UK gliding accidents from 2002 - 2006 deemed to have pilot-related causes. A second rater was also used to ensure reliability of the ratings. Analysis was performed according to two levels of pilot experience (10 hours or fewer experience as pilot in command (PIC) and more than 10 hours PIC), in line with the research findings by AOPA (2006) and those of Jarvis and Harris (2007b), showing that 10 hours PIC or fewer is a particularly vulnerable experience bracket. Factors such as injury level, damage, pilot age and wind strength were also analyzed to see if systematic effects were evident between the two experience groups.

Data

Prior to the present study it was anticipated that the relatively limited pool of UK gliding accidents involving low-hours pilots would be broken down further. Therefore in order to enhance the reliability and validity of the analysis two further years of accident data were added to that used in the previous study (Jarvis & Harris 2007b), meaning that a five year period was used (2002 - 2006).

The British Gliding Association (BGA) database of all UK gliding accidents and incidents over five years from 2002 to 2006 was used (source: British Gliding Association 2007b). This database contained details including pilot age (years); total experience in command (hours); aircraft type; severity of injuries; damage to the glider; and a narrative description of what happened. In line with the previous study (see Chapter 2) accidents resulting in no injury or damage were dropped from the analysis along with ground-handling accidents (such as towing out winch cables or pulling gliders out of hangers), and those without necessary information about the pilot's experience. Where possible the short accident descriptions contained within the BGA database were supplemented with segments from AAIB or BGA reports, as well as additional information from other fields in the database (with the strict exception of pilot experience for the purposes of coding). Hence for each accident a narrative was produced that, in a number of cases, contained more information than the original BGA database text description.

Three categories containing numerical or ordinal data in the BGA accident database were used during statistical analysis. These were:

Nature of Injury: The British Gliding Association categorises accidents into injury categories: 'none', 'minor', 'serious', 'fatal' (British Gliding Association 2007). All accidents contain an entry in this field.

Damage: All accidents in the database contain an entry of either 'none', 'minor', 'substantial' or 'write-off'.

Wind strength: Just over 80% of accidents contained numerical detail on wind strength (in knots) while several more contained text. The rest were dropped from the analysis relating to wind strength. Many of those containing numerical data were in the form of a range (e.g.15 – 20 knots). In these cases the median was used as the value for the purpose of analysis (e.g. 17.5 knots). Where a sign was used (i.e. 10 + knots) only the numerical value given was counted (e.g. 10 knots).

Identification of pilot-related accidents

All accidents from the BGA database (2002 – 2006) were categorized into either primarily 'Pilot-Related' or 'Other' cause ('Technical', 'External' or 'Unknown') using guidelines drawn up and agreed by a group consisting of three subject matter experts and an aviation human factors professional. All members of this group were experienced instructors on gliders and general aviation aircraft. The guidelines are shown in Appendix A. It was pointed out to the raters that 'pilot-related' did not necessarily mean pilot culpability, responsibility or blame, but simply that the accident could be reasonably attributed to the actions taken by the pilot without an overwhelming and unavoidable influence of external or technical factors.

The definition of ‘Pilot-Related’ cause was based upon Hollnagels’ definition of human error (Hollnagel 1998). To be defined as a ‘Pilot-Related’ cause there had to be an identifiable performance shortfall in terms of the actions (or inactions) on the part of the pilot in command together with a reasonable opportunity for the pilot to act in such a way that could have avoided the accident. If a ‘Technical’ or ‘External’ factor was identified as being the seminal accident event then the accident was considered as non-pilot related (i.e. placed in the ‘Other’ category).

The guidelines for coding an accident as being the result of a ‘Technical Factor’ were that the aircraft would have been deemed unserviceable had the failure been apparent before flight. If a ‘Technical Factor’ was induced by the abnormal operation of the glider (outside its operating limitations) this was deemed to be pilot induced. An ‘External Factor’ was regarded as any reasonably unforeseeable and/or unavoidable factor external to the glider that made the flight difficult beyond the skills that could reasonably be expected of a competent pilot. Where ‘External Factors’ were deemed to have been reasonably avoidable or foreseeable then they were again deemed to be ‘Pilot-Related’. Furthermore, difficult flying conditions were only counted as ‘External Factors’ where there were no reasonable signs or expectation of such conditions occurring. A lack of rising air (thermal, wave or ridge lift) was not regarded as an ‘External Factor’ since such ‘lift’ is not reliable and it is also not required for safe glider operation. If it was not possible to identify positively any ‘Pilot-Related’, ‘Technical’ or ‘External Factor’ (i.e. where no causal events could be determined by the rater) the accident cause was categorized as being ‘Unknown’.

The pilot-induced category was further broken down into accidents which involved solo glider pilots and those that involved gliders with more than one person aboard (whether the second person was a passenger or a crew member). Within the latter group, a further designation was made for accidents occurring during training flights (where a pre-solo, or inexperienced pilot was, or had been, handling the glider under the supervision of an instructor pilot). In such flights the instructor is designated as P1 (the pilot in charge) and the student as P2 (the second pilot), even though the student would normally be doing most of the handling. Only those instructional flights where the P2 actions could have influenced the flight in any way were included (and termed 'genuine instructional flights'). This meant that trial gliding flights, gift flights, or flights where the student pilot had no input were not included in the category. Legally, accidents occurring on a training flight would always be the responsibility of the instructor, even though some causality could be due to the actions of the second pilot or trainee (P2). Accident data reflect this by attributing these accidents to P1, meaning they will always be categorised in the 'experienced pilots' group, regardless of the circumstances.

Reliability of the ratings

Since the variable of interest was pilot experience, this was deleted from all reports prior to any categorisation so that all ratings were performed 'blind' thereby avoiding any associated bias on the part of the raters.

In accordance with previous research using large samples of accident data (e.g. Gaur 2005), to establish reliability a random sample of accidents was independently categorized by the primary investigator and an independent rater at each stage of the rating process. Two tests were performed at each stage. In order to check observer consistency a sub-sample of accidents was re-categorized by the primary investigator two weeks after initial rating to establish the intra-rater reliability (a factor omitted in many studies). Such an intra-rater reliability (observer consistency) test is used to check that the researcher's method of coding had been consistent across the database. After this, an inter-rater reliability test was performed by an independent rater to be sure that the researcher had not coded in an idiosyncratic way (Robson 2002). The independent rater was an experienced pilot of gliders and commercial aircraft as well as a crew resource management (CRM) instructor with training in human factors. A random sample of 139 accidents was used (over 25%) to check the reliability of the 'pilot related' designation. A sample of only 100 accidents were used to check the solo/dual and 'genuine instructional flight' designations, but this still represented over 25%

because these designations were only performed on the pool of accidents already designated as pilot-related. Comparisons between the ratings were performed using cross-tabulations and a value of Cohen's Kappa was produced for each test. Robson (2002) cites a commonly used rule for assessing the observer agreement from a calculated Cohen's Kappa co-efficient, as first proposed by Fliess (1981).

Kappa of 0.40 – 0.60: 'Fair'
Kappa of 0.60 – 0.75: 'Good'
Kappa of above 0.75: 'Excellent'

The results for all measures of inter-rater and intra-rater reliability were 'good' or 'excellent' (Table 4.1).

Table 4.1. Reliability figures for all stages of reliability ratings

	Number in sample	Inter Rater Reliability		Intra Rater Reliability	
		% Agreement between raters	Cohen's Kappa	% Agreement between raters	Cohen's Kappa
Pilot related / Other?	139	95%	0.87	96%	0.88
Solo / Dual	100	100%	1	100%	1
Genuine training flight?	100	95%	0.64	98%	0.85

Data analysis

Initially, Fisher's exact tests were used to establish whether significant differences existed between the inexperienced pilots (10 hours or fewer PIC) and experienced pilots (over 10 hours) in terms of pilot related accidents, as well as other factors from the accident database such as injury severity and aircraft damage.

Wind strength data were compared between the two experience groups. The mean strength of wind for accidents occurring to pilots with 10 hours and under was 8.25 knots (sd = 4.1, SE = 1.097), and to over 10 hours pilots it was 10.3 knots (sd = 5.51, SE = 0.428). An independent measures t test on this data showed that the difference between the two means was non-significant ($t = -1.36$, $df = 178$, $p = 0.174$, two tailed). This indicated that the difference in mean wind strength could have occurred by chance, and that wind strength was less likely to be a confounding variable in the differences between the two groups in terms of accident rates.

Results

Of 469 accidents, no causal factors could be determined for 19 (4%), hence these were eliminated from the analysis. Of the remaining 450 accidents, 418 occurred to ‘over 10 hours’ pilots, of which 331 were deemed to have been ‘pilot-related’ (79%). For pilots with 10 hours PIC or fewer there were 32 accidents, 28 of which were designated ‘pilot related’ (88%). A Fisher’s exact test on these data (shown in tabulated form in Table 4.2) was non-significant (p=0.36, two-tailed) suggesting that the distribution of causes (‘Pilot-Related’ or ‘Technical or External’) was randomly distributed between pilots across the two levels of experience.

Table 4.2. Numbers of pilot-related accidents split by experience group.

	Pilot-induced	Other (Technical or external)	Total
10 hrs or fewer	28	4	**32**
Over 10 hrs	331	87	**418**
Total	**359**	**91**	**450**

The mean age of pilots in the 10 hours and under group was 47.2 yrs (sd = 15.76). For the over 10 group it was 51.72 years (sd = 14.2). Pilots with 10 hours or fewer would be expected to be slightly younger than those with more than 10 hours on average (since age must necessarily increase with experience!). Because of this a one-tailed t-test was run on this data and showed that the difference was in fact not significant (t = -1.537, df = 329, p = 1.125).

Table 4.3 shows the number of pilot-related accidents leading to injury and damage for the two pilot experience groups. Fisher’s exact tests on these data show no significant association between the degree of injury and the experience group over the five years being studied (p=0.701). The same is true of aircraft damage analyzed by experience group (p=0.272). There is therefore no evidence to suggest that the accidents suffered by inexperienced pilots were different in terms of their consequences to those suffered by more experienced pilots.

Table 4.3. Total pilot-related accident numbers for the two experience groups broken down by injury severity and damage classification.

		Under 10	Over 10	Total
Injury	Fatal	0	9	9
	Serious	0	18	18
	Minor	3	40	43
	None	25	264	289
Damage	Write off	0	36	36
	Substantial	11	107	118
	Minor	17	186	203
	None	0	2	2

It was found that 70 pilot-related accidents involved gliders with two people on board (coded as 'dual'). Of these, 34 were coded as 'genuine instructional flights' (where the 'P2' had some input under instruction). It was important to identify this small subset of pilot-related accidents at this stage so that it could be analysed further in subsequent studies, alongside the low-hours group. If differences were then found between instructional accidents and accidents occurring to low-hours pilots then this would indicate important issues in terms of what occurs when pilots fly solo. No analysis could be done on these accidents at this stage because they all fell into the 'over 10 hours' pilot bracket (by definition, due to having an instructor on board the glider).

Discussion

The finding that 80% of accidents overall were 'pilot related' is only slightly higher than previous findings across the whole aviation domain (O'Hare 1994, BASI 1996, AOPA 2006), and is almost identical to the figure from US gliding activity reported by van Doorn and de Voogt (2007). Although the low-hours group had a higher percentage of pilot related accidents as predicted (88%) it was not statistically significant in comparison to other pilots. However this does show that the high accident rate among low-hours pilots is due to pilot-related causes rather than other systematic variables. This is further supported by results showing that no statistically significant difference existed between the two experience groups in terms of wind conditions and pilot age.

Analysis of the primary causal categories demonstrates that accidents involving pilots with 10 hours flight experience or fewer were not statistically different in their distribution of injury and aircraft damage to those of more experienced pilots. This means that statistically, low-hours pilots were just as likely to damage a glider or injure themselves in an accident as other pilots. Figure 4.3 shows that no serious injuries occurred to low-hours pilots in the period 2002 and 2006. However the statistics indicate that this was likely to be due to the relatively low number of flights accrued by this group, as opposed to a systematic cause. Low hours pilots had a marginally higher percentage of accidents leading to substantially damaged aircraft (39% of all accidents, as opposed to 32% for other pilots), and it has been shown that there is a high correlation between damage and injury in gliding (van Doorn & de Voogt 2007).

This study identified that out of 469 UK gliding accidents occurring during the period 2002 - 2006, three-hundred and fifty nine (359) were caused primarily by the pilot (see Table 4.2). Thirty-two of these occurred to pilots with 10 hours or fewer as pilot-in-command, and 34 occurred to student pilots under instruction.

The findings of the present study, along with the previous study (Jarvis & Harris 2007b) provide robust evidence from which to conclude that low hours solo glider pilots in the UK are relatively vulnerable to accidents, and that those accidents are somehow linked to their piloting of the glider, as opposed to external or technical influences. Clearly, more research is needed to look into pilot-related accidents in order to investigate the accident causes, particularly between low hours and more experienced pilots. To this end the 359 accidents identified in this study, along with their narrative descriptions and some supplementary material, were prepared for use in subsequent studies aimed at identifying deeper causal accident factors.

Chapter 5 - Flight Phase Accident Analysis

This research was published (along with chapter four) as:

Jarvis, S. Harris, D. (2008), *Investigation into Accident Initiation Events by Flight Phase, for Highly Inexperienced Glider Pilots*. International Journal of Applied Aviation Studies Vol 8, Number 2. FAA Academy.

Introduction

It has been shown that low-hours UK glider pilots (defined as under 10 hours total time as pilot-in-command) have a much higher accident rate than their more experienced counterparts (Jarvis and Harris 2007b). It has also been established that the vast majority of UK gliding accidents, including those occurring to low hours pilots, are pilot-related (See Chapter 4) as opposed to being caused by other factors (mechanical, external, etc). The majority of accidents contributing to the disproportionate accident rate for low-hours pilots were therefore caused by pilot-factors. It was shown that no-statistical difference exists in the categories of injury suffered by low-hours pilots and hence it is clear that low hours UK glider pilots are less safe than more experienced pilots, and that the reason is somehow related to the pilots' general levels of skill.

In order to begin to investigate the reasons behind this it was necessary to analyse where accident events were most likely to occur during flights. It was possible that some parts of the flights were causing more problems than others, particular for low-hours pilots, or that some specific tasks in the flight were more likely to trigger mistakes or errors caused by pilot inexperience. Such an analysis could be useful in directing training interventions for pre-solo pilots, and interventions for low-hours post solo pilots. Additionally any subsequent overall human factors analysis on the accidents would be of greater use if it could be determined at which points in the flights human factors problems emerged.

A study of 143 US gliding accidents by van Doorn and de Voogt (2007) found that based simply on a frequency analysis, over half of all accidents (52.4%) occurred on landing, while 30% occurred during the ‘cruise’. Most fatal accidents in gliders took place during the cruise (van Doorn & Zijlstra 2006); 36% of these accidents were found to end in fatality, whereas only 10% of landing accidents did so (van Doorn & de Voogt 2007). These studies did not, however, break down the accident statistics by pilot experience, nor did they take into account flight exposure. Furthermore they could be criticized as lacking explanatory power as a result of the nature and extent of the phases used. Van Doorn and de Voogt (2007) used only four phases to describe all accidents; assembly, tow, cruise and landing. Although problems can occur during ‘assembly’, it is problematic to compare this numerically with other phases of flight since many gliders are kept in hangers or flown many times per assembly, hence many glider flights do not include this phase at all. This leaves only three in-flight phases, all of which include numerous flight components (see British Gliding Association 2003, Stewart 1994, Piggott 1997) meaning that categorization using such a taxonomy would be questionable in terms of its utility in identifying specific problem areas. For example, an accident deemed to have occurred in the ‘cruise’ phase could have taken place in midair (e.g. a collision or overstress leading to break-up), in the circuit or approach to a field landing, during an attempted out-field landing or as a result of unintentional ground impact. There is no current evidence to suggest differences may exist between experienced and inexperienced glider pilots in terms of accident flight phase, and no such research has been carried out.

Studies of powered aeroplane accidents tend to use a greater number of flight phases. Many studies report accidents by the same flight phase descriptions as used by the NTSB. Seventeen NTSB flight phase categories exist; ‘standing’, ‘taxi’, ‘takeoff’, ‘climb’, ‘cruise’, ‘hover’, ‘descent’, ‘descent - emergency’, ‘descent - uncontrolled’, ‘approach’, ‘landing’, ‘emergency landing’, ‘go around’, ‘missed approach’, ‘manoeuvring’, ‘other’ and ‘unknown’ (FAA 2008). Unfortunately most of these are not applicable to gliders which have much less clearly defined phases of flight. Phases such as ‘taxi’, ‘go around’ and ‘missed approach’ would clearly not apply to gliders and the ‘cruise’ phase would not easily fit into the operation of a glider in the way it does a powered aeroplane, since an engine is required to sustain cruise flight. ‘Manoeuvring’ is an example of a flight phase that could be very applicable to gliding accidents; it is known that serious and often fatal accidents do occur from low turns (Jarvis 2004, Jarvis and Harris 2007a). The 'descent' and 'climb' phases could apply to almost all of a glider's flight, since all the time that the glider is not climbing in lift it is descending in a glide, but the glider pilot does not have the same degree of control over such phases as does a pilot of a powered aircraft. An analysis with a more detailed breakdown of flight phases allowing more explanatory power and hence better-targeted remedial interventions is required to encompass the unique operation of a glider.

In general aviation, manoeuvring and landing are consistently cited as being the most common accident flight phases. Again, based solely on frequency counts, it has been shown that 39% of accidents occur during the landing phase (O'Hare 1994). AOPA (2006) reported that 38.9% of non-fatal accidents happened during 'landing', more than any other flight phase, while observing that 'manoeuvring' was the most common flight phase for fatal accidents (22.8%). In a study of insurance claims, Lenné and Ashby (2006) reported that the landing and taxiing phases accounted for 55% of all non-fatal general aviation accidents in Australia.

The only study to specifically analyse accidents for highly inexperienced general aviation pilots used NTSB flight phase descriptors to show that US solo student pilots (i.e. very low hours solo) had 146 (45%) accidents during landing over a two year period (Baker et al 1996). Very few of these however led to injuries. Touch-and-goes (a manoeuvre including a landing and take-off) accounted for 78 (24%) accidents. Some phases were notably under represented; only four (1%) occurred during turns, seven (2%) occurred in the circuit (downwind or base-leg) and eight accidents (2%) occurred during the final approach to land. This suggests that a considerable section of the flight (from entering the circuit until landing) appears to be a relatively safe phase of flight for very inexperienced general aviation pilots, accounting for only 4% of accidents to early solo pilots.

The use of flight phase categorisation in accident descriptions is far from straightforward. While the accident reports of many national investigation organisations contain flight phase categorisations, many, such as the UK's Air Accident Investigation

Branch, do not. It is often unclear how a flight phase descriptor relates to the accident, or how it has been defined. Several definitions are often possible, leading to a different selection of phase. For example the 'accident flight phase' may be defined as the stage causing the damage/injury, the phase containing the initial error, the emergency phase (AOPA 2006) the phase containing the most obvious departure from safety, etc. Furthermore, accidents are often the result of a chain of events rather than a single event or error (Wiegmann et al 2005). Therefore taxonomies and classifications (including databases such as the NTSB) that categorise by a single flight phase may risk over-simplification, as a series of causal events may have accumulated during the flight. In particular there is a danger that the flight phase in which the initiating event took place is not categorised. This scenario is particularly likely in gliding as a result of the difficulty in regaining lost energy in terms of height and/or speed. For example: misjudgment of height when entering the circuit may lead to poor positioning of the base leg with little energy to reach the airfield, subsequently resulting in a slow approach and heavy landing. In such a case the accident 'flight phase' might be categorized as the stage in which the damage/injury was sustained (landing) thereby failing to identify that the initiating event occurred much earlier in the flight. Recognizing the issue of multiple events Wiegmann et al (2005) categorized accidents using any number of flight phases but labeled only one of these to be the 'seminal' phase, in which the initiating event was deemed to have taken place. This same approach was initially used in the study of North Sea Helicopter Safety (Ingstad et al 1990). Lenne & Ashby (2006) also used a similar method by identifying the first crash occurrence noted in the accident narrative. Baker et al (1996) used a similar method. It was pointed out that the flight phase allocated did not always coincide with the NTSB phase for the same accident because it was based upon "the period when the problem arose" (Baker et al 1996).

Although research dedicated to accident flight phase has been conducted for general aviation (and a small amount for gliding), it has been shown that there has been little attention to the relationship between pilot experience and the flight phases of initiating accident events. Furthermore while some research has used frequency data and accident totals (BASI 1996, AOPA 2006, van Doorn & Zijlstra 2006; van Doorn & de Voogt

2007) no research has provided comparable accidents rates, possibly due to the difficulties of obtaining exposure data for pilots of differing experience levels.

This study produces a detailed flight phase template applicable to glider operations, and using UK gliding accident data, compares the flight phases in which the initiating event preceding an accident occurred with respect to highly inexperienced pilots (10 hours or fewer experience as pilot in command) and more experienced pilots (over 10 hours). Furthermore, the data obtained are used to provide estimates of accident rates (both in terms of hours flown and number of launches) for both phase of flight and pilot experience.

Method

All accidents from 2002 - 2006 deemed to be 'pilot-related' from the previous analysis (see Chapter 4) were used in this study. The same categorization and narrative descriptions from the UK accident database were used (source: British Gliding Association 2007b). The analysis of accident data progressed in two stages. Firstly, the numbers of pilot-related events were identified within each accident report with the 'seminal event' being categorized as the first to occur. Secondly, a detailed flight phase template was constructed using subject matter experts and task analysis techniques in order to conduct a flight phase analysis on the seminal events. All accidents were analyzed according to two levels of pilot experience (10 hours or fewer experience as pilot in command (PIC) and more than 10 hours PIC), in line with the previous research findings. Since pilot experience level was a key variable of concern in the research, this information was removed from the accident descriptions during categorization to avoid influencing the process, in common with the previous study.

Stage 1: Identification of pilot-related contributory events

Within the BGA accident database, it was clear from performing the analysis during the previous study (Chapter 4) that there were many cases where a number of events occurred, usually in different flight phases, in order to cause the accident. The individual major 'pilot contributory' events in each accident were identified from the accident narratives using guidelines in order to increase consistency (see Appendix B). The guidelines were drawn up by the investigator and agreed by the same three subject matter experts used previously. Some of the guidelines written previously (shown in Appendix A) were relevant and therefore included. Each accident report was categorised by any number of events, in the order that they occurred. Events occurring within the same flight phase were designated as separate events. Following this, the seminal event was identified; this being defined as the first event in the sequence (cf. Ingstad et al 1990, Wiegmann et al 2005). Reliability tests were performed in the same way as for the other stages.

Stage 2: Flight phase analysis

It has been shown that many of the NTSB flight phase descriptions are not fully applicable to gliding, and that no adequate flight phase designators have been used to research gliding accidents. Therefore a framework for glider flight phases was prepared that supplemented that used by van Doorn and de Voogt (2007) that categorised flights into three phases; tow, cruise and landing.

A high-level mission analysis utilizing concepts drawn from process charting methods (see Kirwan & Ainsworth 1992) was undertaken to breakdown the operation of a glider into meaningful, quasi-independent flight phases. Resources such as the BGA instructors' manual (2nd ed, British Gliding Association 2003), Piggott (1997) and Stewart (1994) were used in this process along with a number of subject matter experts (experienced gliding instructors). This analysis was performed to produce a two-level flight phase template. The resulting template consisted of 25 flight phases in total, grouped within six higher-order phases (pre-flight; launch; in-flight phase; circuit; approach and landing). Agreement was reached between the subject matter experts that the final template was representative of all aspects of UK gliding operations. This coding template is shown in Figure 5.1.

The flight phase analysis highlighted the requirement to separate accidents occurring during attempts to land at an airfield from accidents occurring while attempting to land in an unfamiliar field, which can often occur when insufficient lift is found to continue the flight. Off-airfield landings (also known as field landings) are common in gliding but involve unique tasks such as assessing field size and suitability, and positioning a circuit to an unfamiliar site with no primary height information. This is accepted as common practice in gliding (rather than an emergency) particularly when a pilot is attempting a cross-country soaring flight. It was therefore necessary to be able to identify such accidents during analysis in case they had a substantial effect on the findings. Therefore prior to attaching the flight phase descriptors, each accident was classified by its location ('airfield' or 'off airfield'). Accidents in the circuit or approach phase of the base airfield (or intended landing airfield) were labeled as 'airfield' accidents, whereas those occurring outside the circuit pattern of the airfield were treated as ‘off-airfield’ accidents. Additionally, accidents occurring during intentional cross-country flights were noted.

Accidents following launch failures required identification for similar reasons. Launch failures can require unique maneuvers such as regaining flying speed at low altitude and flying low abbreviated circuits. Therefore it was essential to have the ability to separate these from normally launched glider flights for the purpose of analysis. A failed launch recovery was deemed to be any accident flight where the launch failure was the seminal event and from which the glider was unable to join the normal circuit from the high key position, without encountering lift. This included: breaks (weak link/ cable break/ rope break/ unexpected release) as well as power failures (winch power failure, tug wave off / power loss). Where the main causal factor led to an aborted or failed launch (e.g. ground loop, over rotation) then this was not coded as a failed launch recovery, since in such a case the launch failure was not the seminal event in the accident chain, but a consequence of another event.

All accident events were categorized using the flight phase template (figure 5.1). In addition, each accident was further categorized as being normal launch/launch failure; airfield/off-airfield.

1. **Pre-flight, after boarding**

2. **Launch**
    a. Ground run [Only if attached to cable]
    b. Pre-rotation initial climb (Winch)/ Airborne, pre-climb (Aerotow)/PIOs
    c. Rotation and establishing main climb (Winch)/Initial Climb (Aerotow)
    d. Established Climb
    e. Aerotow cruise (cross-country towing etc)
    f. Release
    g. Recovery to normal flying speed, period prior to manoeuvre/approach
    h. Non normal aerotow: Low tow, 'Boxing the tug'

3. **Flight phase**
    a. General flying (practicing manoeuvring, local soaring/flying etc)
    b. Serious soaring (circling, dolphin, eights, street flying, cloud climb, leaving, joining etc)
    c. Search/descent (Usually during cross-country/ extended soaring. Search for lift / search and inspection of field to land, prior to circuit/ abbreviated circuit commitment)
    d. Ridge soaring
    e. Wave flying
    f. Final glide (incl. comp finish/ return to airfield/ 'stretching the glide')
    g. Aerobatics/ intentional unusual manoeuvring (stall/spin/steep turns)
    h. Immediate evasive airborne manoeuvre to avoid imminent collision

4. **Circuit**
    a. Circuit join. [Phase between 3 & circuit to land] (including prep, wheel down, decisions on circuit/landing direction). NOT choice of field itself, that is 3c (search/cruise)
    b. Circuit (from high key to final turn, joined from anywhere) Include as seminal where accident descriptions begin from a poorly positioned final turn or final approach from turn (too high, low, far, close)
    c. Abbreviated circuit or non-standard manoeuvring to land
    d. Final turn (from normal or abbreviated circuit only)

5. **Approach**
    a. Approach after circuit or flying (approach other than 4b)
    b. Approach as straight ahead recovery from launch failure

6. **Landing/Ground**
    a. Flare/ Hold off. Includes ballooning and PIOs
    b. Ground run [post landing or after ground cable release]

Figure 5.1. Flight phase categorization template

Reliability of the ratings

In common with previous studies, information about pilot experience was deleted from all accident reports prior to any categorisation. The techniques used for reliability testing were exactly the same as used in chapter four. Inter-rater and intra-rater tests were performed on a random sample of just over 25% of the data (100 cases). Statistical comparisons were performed using cross-tabulations.

For stage 1, accident events were identified, and 'number of events per accident' was used as the variable to be compared between raters. In stage two the seminal events for each of the 100 accidents in the reliability sample were highlighted and coded using the flight phase template. The category chosen for each seminal event was compared between raters.

The results for all measures of inter-rater and intra-rater reliability were 'good' or 'excellent' at all stages (Table 5.1). The classification matrices for 'number of events' and 'high level flight phases' are also shown (Tables 5.2, 5.3, 5.4 and 5.5).

Table 5.1. Reliability figures for all stages of reliability ratings

	Number in sample	Inter Rater Reliability		Intra Rater Reliability	
		% Agreement between raters	Cohen's Kappa	% Agreement between raters	Cohen's Kappa
Launch Failure or not?	100	99%	0.95	98%	0.91
Airfield / Off airfield	100	98%	0.95	99%	0.98
Intentional cross country flight?	100	98%	0.95	99%	0.97
Number of Events per accident	100	87%	0.61	89%	0.67
High Level Flight Phase categorization	100	87%	0.84	96%	0.95
Low level Flight Phase categorization	100	81%	0.79	91%	0.90

Table 5.2. Intra-rater classification matrix for establishing reliability In terms of the number of events designated to each accident.

INTRA RATER		Original Rater			
		1 event	**2 events**	**3 events**	**total**
Second Rater	**1 event**	58	3	0	61
	2 events	5	21	2	28
	3 events	0	3	8	11
total:		63	27	10	100

Table 5.3. Intra-rater classification matrix for establishing reliability In terms of the number of events designated to each accident.

INTER RATER		Original Rater			
		1 event	**2 events**	**3 events**	**total**
Second Rater	**1 event**	59	2	0	61
	2 events	2	23	4	29
	3 events	0	3	7	10
total:		61	28	11	100

Results

Results - Treatment of Data

The 359 accidents identified as being pilot-related were subjected to the main flight phase analysis. Since accident causation was of primary concern in the research, each accident was classified according to its seminal event (as identified in stage 1). For example, an accident designated as occurring in the launch was one in which the seminal event took place whilst the glider was launching. After classification of events, those other than the seminal events were dropped from the analysis.

The distribution of seminal events occurring during the various flight phases with respect to pilot experience groups were initially analyzed using Fisher's exact tests. Odds ratios with associated confidence intervals were then calculated between the two experience groups for all six top level flight phases.

Although such analyses can be used to compare the frequency of accidents in one group with that of the other and identify where accident features were disproportionately distributed between groups, they cannot account for differences in flying exposure between the groups (including flying that did not result in an accident). For this, a measure of exposure was required in order to produce accident rates for comparison.

The exposure estimate from Jarvis & Harris (2007b) required re-calculation due to the additional two years of accident data needed to cover the five-year period 2002 - 2006. The exposure estimates were recalculated using the original method from Jarvis and Harris (2007b) but including two additional years of data collected from 2002 and 2003 in order to cover the whole period 2002 to 2006 inclusive. The 10 hours (and fewer) exposure estimate was subtracted from BGA annual totals to provide data for the two groups; pilots with 10 hours or fewer and those with more than 10 hours as pilot in command. On this basis the estimated total number of launches taken from 2002 – 2006 by pilots with 10 hours PIC or fewer was 29,924 with an upper 95% confidence boundary of 35,301 launches and a lower 95% confidence boundary of 24,548 launches.

The estimated total number of hours flown was 11,553 hours (upper 95% confidence boundary of 14,017 hours and a lower 95% confidence boundary of 9,089 hours. The mean calculation of flying exposure by pilots with over 10 hours PIC during the same period was 1,609,810 launches and 696,041 hours.

Table 5.4. Classification matrix for intra rater reliability ratings of the six high level flight phases (FP1 - FP6).

INTRA RATER RELIABILITY	Original Rater						
Second Rater	FP1	FP2	FP3	FP4	FP5	FP6	
FP1 - Pre-flight	6		1				**7**
FP2 - Launch		9		1			**10**
FP3 - General Flying			28				**28**
FP4 - Circuit	1			21			**22**
FP5 - Approach			1		14		**15**
FP6 - Landing						18	**18**
	7	**9**	**30**	**22**	**14**	**18**	**100**

Table 5.5. Classification matrix for inter-rater reliability ratings of the six high level flight phases (FP1 - FP6).

INTER RATER RELIABILITY	Original Rater						
Second Rater	FP1	FP2	FP3	FP4	FP5	FP6	
FP1 - Pre-flight	5						**5**
FP2 - Launch	2	8		1		1	**12**
FP3 - General Flying			27				**27**
FP4 - Circuit			3	19	2		**24**
FP5 - Approach		1		2	12	1	**16**
FP6 - Landing						16	**16**
	7	**9**	**30**	**22**	**14**	**18**	**100**

Results - Stage 1: Accident events.

Within the 359 'pilot-related' accidents classified in the previous study (Chapter 4), 545 causal events were identified in total. Only three accidents were categorized as containing four events, and so these were combined with the three-event accident group, to make up a group labeled 'three or more events' (see Table 5.6). The resulting analysis using a Fisher's exact test showed a significant association between pilot experience group and the number of events in the accident sequence (p=0.016). Further analysis of standardized residuals indicated that the '10 hours and under' group had a significantly higher proportion of accidents where three or more events were identified in the analysis (standardized residual of 2.3, p=0.016).

Table 5.6. Numbers of events identified in the accident narratives, split between the two experience groups.

		Under 10	Over 10	Total
No of events	**1 event**	11	203	214
	2 events	10	97	107
	3 or more events	7	31	38

Results - Stage 2: Flight phase analysis

Launch failures

Twenty-four pilot-related accidents happened after launch failures. Of these 23 occurred to pilots with over 10 hours of experience and one to a pilot in the 10 hours of fewer group. A Fisher's exact test on this data gave a two-tailed result of p=0.707, showing that a systematic effect was statistically improbable.

However, subsequent analysis found that nine out of 23 of these accidents occurring to the experienced pilots were associated with training flights (where the experienced pilot was the instructor pilot, but the handling pilot was a trainee). Indeed, from the accident narratives it can be established that all of these accidents were 'simulated' launch failures (where the instructor deliberately releases the launch cable from the glider prematurely as if the cable or weak-link had broken, in order to test whether the student can deal effectively with the 'emergency'). All of these occurred during winch launching. None of the other launch failure accidents were of the simulated type (understandably, since they were mostly solo pilots). No data exist to show how many simulated exercises are performed compared to the number of real launch failures, and no such records are kept. Without such exposure measures, it is impossible to draw conclusions about how likely accidents are to occur after simulated launch failures as opposed to real ones. However the data do show that simulated launch failures in training cause nearly as many accidents as real launch failures.

Fine grained flight-phase analysis showed that of the nine simulated launch failures leading to accidents, five of them had seminal events occurring in the recovery to flying speed after the failure had occurred (phase '2g' in the template, Figure 5.1). Three caused serious injuries to one or more of the occupants, and two had no injuries associated with them. No other fine-grained phase was associated with more than one launch failure accident.

Off airfield accidents

Table 5.7 shows that of the 359 pilot-related accidents, 113 were identified as occurring off the airfield (meaning outside the circuit). All but one of these involved pilots with over 10 hours flying experience. The result of the Fisher's exact analysis suggested that accidents occurring away from the base airfield were not randomly distributed across the experienced and inexperienced groups ($p<0.000$, two tailed). In terms of simple frequency of occurrence, experienced pilots were much more likely to have an accident away from their home airfield. The odds ratio was 13.807, suggesting far greater odds of experienced pilots sustaining this type of accident, although the 95% confidence interval was extremely wide (1.85 – 102.9).

Table 5.7. Total accident numbers for the two experience groups broken down by accident location

		Under 10	Over 10	Total
Accident location	**Airfield / within circuit**	27	219	246
	Off-airfield	1	112	113

Flight phases overall (all pilots)

For this main flight phase analysis each accident was classified only by its seminal event. From this point, all other accident events were dropped from the analysis.

Table 5.8 shows the total number of accidents in each high level flight phase category. Table 5.9 shows the same information broken down by the lower level (fine grained) flight phases.

Table 5.8 shows that of the six high-level phases 'general flying' included most accidents overall (i.e. where the seminal event leading to accident occurred in the general flying phase). The distribution of injuries (Table 5.11) shows that seminal events in the 'launch' phase and the 'general flying' phase led to the most severe accidents. From finer-grained analysis using the sub-phases (Table 5.9) it can be seen that the ground run phase caused most accidents (although relatively few injuries). Recovery to speed after release from launch was associated with ten accidents (causing four serious injuries). Initial events occurring in the rotation into the climb (on launch) were associated with most fatalities (3 in total). For the general flying phase most injuries occurred during ridge soaring (one fatality, two serious and five minor injuries). The search/descent and final glide stages also had high numbers of injuries (three serious and four minor).

Table 5.8. Seminal accident event totals for each high-level flight phase descriptor, as well as a breakdown between the two experience groups. 'Instructional flight' totals are the subset of those seminal accident events occurring during training flights.

Seminal events leading to accidents	All seminal events	UNDER 10 hours pilots (seminal events)	OVER 10 hours pilots (seminal events)	Seminal events during genuine instructional flights
1. Pre-flight	15	3	12	1
2. Launch	41	2	39	9
3. General Flying	94	1	93	2
4. Circuit	71	3	68	3
5. Approach	72	11	61	6
6. Landing	66	8	58	13
Totals	**359**	**28**	**331**	**34**

Table 5.9. Seminal accident event totals for each low-level flight phase descriptor, as well as a breakdown between the two experience groups. 'Instructional flight' totals are the subset of those seminal accident events occurring during training flights.

Seminal events leading to accidents	All seminal events	OVER 10 hours pilots (seminal events)	UNDER 10 hours pilots (seminal events)	Seminal events during instructional flights
1. Pre-flight	15	12	3	1
2a. Ground run	20	19	1	2
2b. Pre-rotation climb	4	4	0	0
2c. Rotation	4	4	0	1
2d. Established climb	1	0	1	0
2e. Aero-tow cruise	1	1	0	0
2f. Release	0	0	0	0
2g. Recovery to speed	10	10	0	5
2h. Non normal aero-tow	1	1	0	1
3a. General flying	12	11	1	1
3b. Serious soaring	10	10	0	0
3c. Search/descent	40	40	0	0
3d. Ridge soaring	15	15	0	1
3e. Wave flying	0	0	0	0
3f. Final glide	16	16	0	0
3g. Unusual manoeuvring	1	1	0	0
3h. Evasive manoeuvre	0	0	0	0
4a. Circuit join.	39	39	0	1
4b. Circuit	23	21	2	1
4c. Abbreviated circuit	6	5	1	1
4d. Final turn	3	3	0	0
5a. Approach (circuit etc)	68	57	11	5
5b. Approach (lnch failure)	4	4	0	1
6a. Flare/ Hold off.	52	45	7	11
6b. Ground run	14	13	1	2
TOTALS	**359**	**331**	**28**	**34**

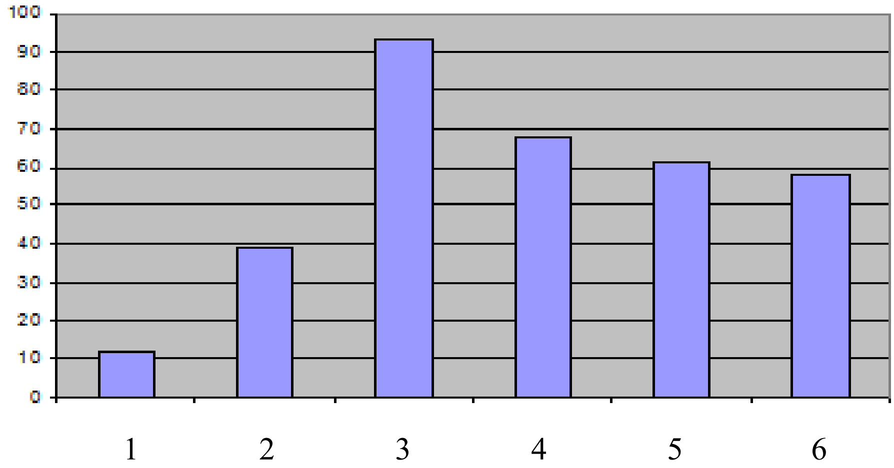


Figure 5.2. Accident totals for pilots with <u>over 10</u> hours PIC, split by flight phase seminal events (flight phases 1 - 6)

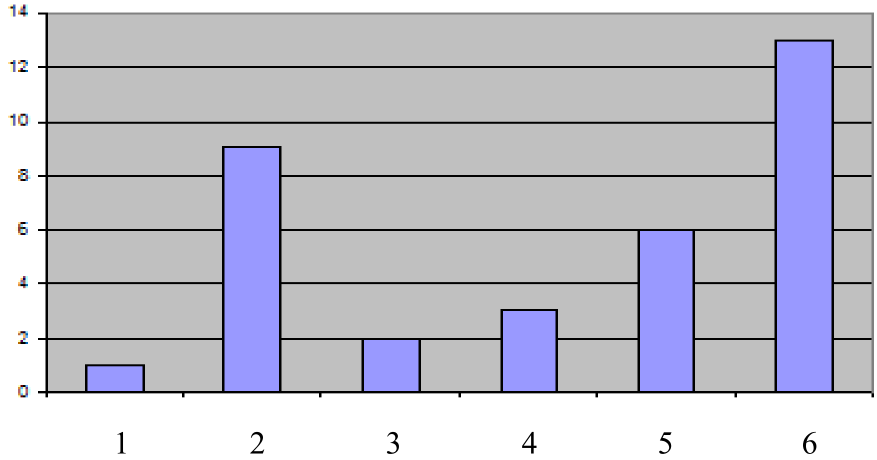


Figure 5.3. Totals for <u>instructional accidents</u> (subset of over 10 hours pilots) split by flight phase seminal events (flight phases 1 - 6)

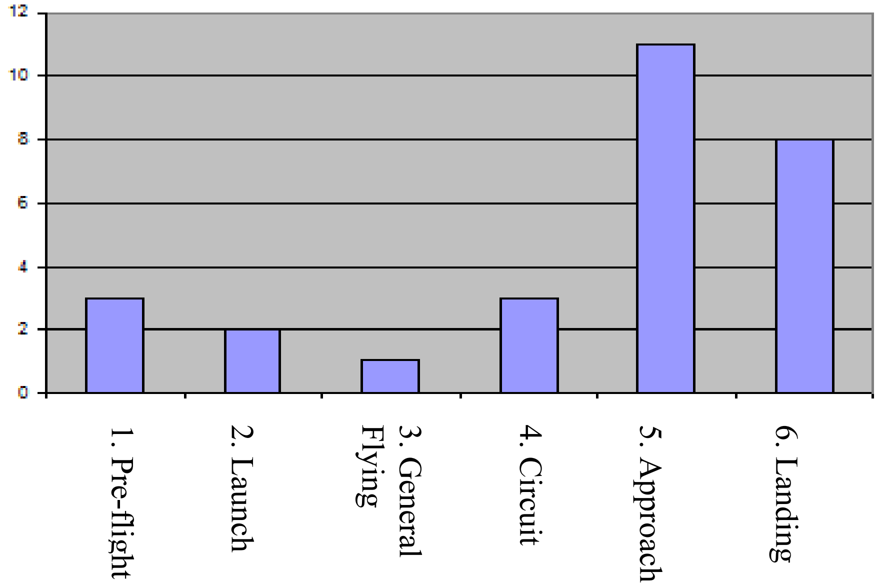


Figure 5.4. Accident totals for pilots with <u>10 hours or fewer</u> PIC, split by flight phase seminal events

Instructional Accidents (Training accidents)

Figures 5.2, 5.3 and 5.4 show flight phase totals displayed for ease of comparison of accident trends across flight phases, for the two groups as well as for the sub-group of training accidents (identified as genuine instructional accidents). Figure 5.3 clearly shows that the distribution of training accidents is closer to the distribution for pilots with 10 hours or fewer than it is to pilots with over 10 hours, despite training accidents being a sub-set of the latter.

In total, there were 34 accidents coded as ‘genuinely instructional’, meaning that 10.27% of accidents occurring to the over-ten pilots happened while they were instructing students (34 out of 331). Figure 5.5 shows how this figure splits between flight phases; in other words the percentage of accidents in each flight phase that happened while instructing. For example, 23% of launch accidents occurring to pilots with over ten hours were accounted for by instructional flying.

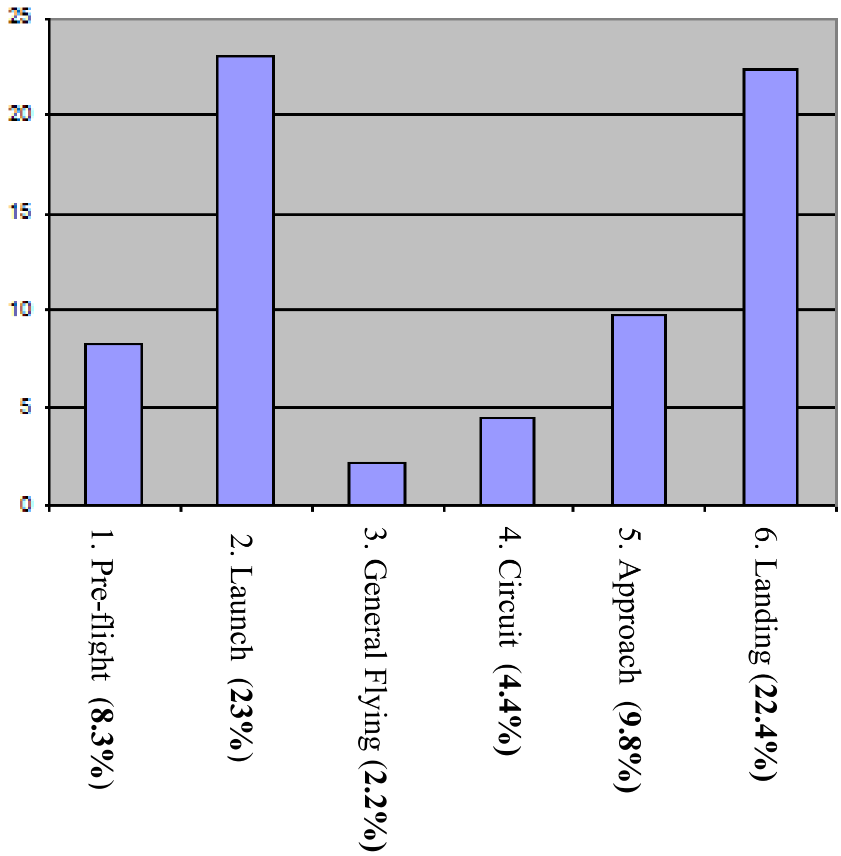


Figure 5.5. The percentage of accidents to pilots with 10 hours or over, attributable to instructional flights, split by flight phase. These flights are a subset of the over-10 hours accidents.

A Pearson's Chi-square test was used to compare the differences between flight phases in terms of instructional and non-instructional flying (within the over 10 pilot group). Table 5.10 shows the data, including the expected cell counts. The result was significant (two-tailed) to a 99.9% level of confidence ($\chi^2 = 25.2$, df = 5, $P<0.001$), confirming that (in terms of flight phase) the pattern of accidents was different for instructional accidents than for non-instructional ones. It can be seen that two cells had expected frequencies of less than 5, which is generally understood to lower the statistical power of the test (Field 2005, Robson 2002). However, most sources point out that while this rule applies to small contingency tables (e.g. 2 x 2) larger tables can sustain a number of such cases, usually deemed to be up to 20% of cells (Field 2005, Coolican 1994). Since in this case less than 20% of cells fall into this category and any associated error would be type two (i.e. finding no effect when one exists), this was deemed acceptable.

Table 5.10. Cross-tabulation used for Pearson's Chi-square test. The population is all accidents occurring to pilots with over 10 hours of experience. This is split by the flight phase in which the seminal event occurred, and whether or not the flight was instructional (as coded). Bracketed numbers are the expected cell counts.

	Non instructional accidents (over 10 group)	*Instructional accidents*	*Total*
Pre-flight	11 (10.8)	1 (1.23)	12
Launch	30 (35)	9 (4.01)	39
General Flying	91 (83.4)	2 (9.55)	93
Circuit	65 (61)	3 (6.98)	68
Approach	55 (54.7)	6 (6.27)	61
Landing	45 (52)	13 (5.96)	58
	297	34	331

Twenty-three percent of all accidents coded as occurring during launch (for the 'over 10 hours' group) happened on instructional training flights (9 out of 39). For landing the figure was 22.4% (13 out of 58). Launch and landing therefore stand out as being the phases that were disproportionately high for instructional flying accidents. Fine grained analysis (Table 5.9) indicates that most instructional training accidents initiated during the launch phase had the seminal event in the 'recovery' to speed after the release (flight phase 2g). Unsurprisingly, all of these were simulated launch failures, including three accidents resulting in serious injuries. Interestingly no such accidents (recovery to speed, ‘2g’) occurred to pilots in the ’10 hours or fewer’ group, so despite causing serious accidents during training, the recovery phase caused few accidents once pilots were solo. It was also notable that few accidents during the launch ‘ground run’ happened on instructional flights, despite these accidents being the largest group of accidents in the launch phase overall. The ‘general flying’ phase and ‘circuit’ phase were associated with very few accidents during instructional flying.

Flight phases split by pilot experience

From figures 5.2 and 5.4, it can be clearly seen that the pattern of accidents across flight phases was quite different between the two pilot experience groups.

Table 5.11. Totals, frequencies and injuries categorized by where the seminal event was deemed to have taken place, in terms of the six higher level flight phases. Launches per accident are calculated from exposure data. The exposure data for pilots with 10 hours or fewer pilot-in-command was derived from the data published in Jarvis & Harris (2007b).

High level Flight Phase	Injuries totals			10 hours or fewer experience			More than 10 hours	
	Fatal	Serious	Minor	Total accidents	Launches per accident	95% Confidence boundary	Total accidents	Launches per accident
1. Pre-flight	1	1	2	3	9,975	± 1,792	12	134,151
2. Launch	4	2	9	2	14,962	± 2,688	39	41,277
3. General flying	2	9	12	1	29,924	± 5,377	93	17,310
4. Circuit	1	0	5	3	9,194	± 1,792	68	23,674
5. Approach	0	3	12	11	2,720	± 489	61	26,390
6. Landing	0	1	4	8	3,741	± 672	58	27,755

For initial statistical comparison of the two pilot experience groups, odds ratios were calculated for accident involvement across the six flight phases. An example of the tabulated data for a single test would be the total number of ‘landing’ accidents against the total number of all other accidents, then all split by the two pilot groups (Example: Table 5.12)

Table 5.12. Example of the tabulated data for the odds ratio calculation (this is for accidents initiated in the landing phase)

	10 hours and under PIC	Over 10 hours PIC	**Total**
Accident initiated in the landing phase	8	58	**66**
Accidents not initiated during the landing phase	20	273	**293**
Total	**28**	**331**	**359**

This was done six times; once for each of the flight phases. Only two phases out of the six resulted in significantly different odds between the two experience groups (i.e. where the 95% confidence interval did not include 1). These were the approach phase (Odds ratio: 2.864, in favour of inexperienced pilots having such accidents, with a 95% confidence interval from 1.28 to 6.42) and the ‘general flying’ phase (Odds ratio: 10.55, in favour of experienced pilots having such accidents, with a 95% confidence interval from 1.41 to 78.765).

However these results only show the odds of one type of accident occurring in relation to other types. They do not show the actual likelihood of these accidents occurring on a given flight. For such analysis, accident rates were required, for both pilot groups, and all flight phases.

Accident rates were calculated for comparison by flight phase between the two pilot groups. In the phases of flight where high numbers of accidents were observed more detailed analysis was performed. The rates across the six high-level flight phases (broken down by pilot experience) are shown in Table 5.11.

Whereas approach and landing were associated with the highest accident rates within the 10 hours or fewer group, the ‘general flying’ and ‘circuit’ phases had the highest accident rates within the experienced pilots’ group. The ‘general flying’ phase was the only phase where the accident rate for the over 10 hours group was higher than the 10 hours or fewer group (estimated at one launch in 17,310 against one in 29,924 for the 10 hours and fewer experience group). Fine-grained analysis of the over 10 hours group showed that element 3c (search/descent) accounted for 40 accidents in the ‘general flying' phase (43% of that category) by far the largest element. No accidents were caused during this phase for the 10 hours or fewer experience group, and further analyses were not undertaken owing to the limited number of accidents involved.

The highest accident rate overall was associated with the inexperienced pilots and having an accident caused during the approach phase (1 in 2,720 launches by the mean estimate and one accident in every 3,209 launches by the lowest estimate). Even the lowest estimate is over eight times that of the more experienced group. The next highest was the landing phase with the lowest estimate of the accident rate at one in 4,413 launches, over six times higher than that for the over 10 hours experience group.

Discussion

Stage one found that 44% of accidents were designated as having more than one ‘pilot induced’ causal event, however the narrative descriptions of accidents involving inexperienced pilots were more likely to describe multiple contributory causes. Therefore it is possible that inexperienced pilots are less able to successfully deal with the consequences of an initial error, and so make further errors (i.e. compound the problem). Equally however this could be due to the quality of the reports and witness statements. It has been shown that inexperienced pilots regularly crash at or close to the airfields (much more so than the experienced group) and this means that the accidents are more likely to be witnessed by other glider pilots. This effect could add detail to the accident narratives, meaning more events were identified.

Direct comparisons of flight phase findings from stage two with previous research are problematic because the definitions of phases used in this research are more detailed than in most previous studies, particularly in gliding. Also this analysis uses seminal events (the first contributory factor leading to an accident – (see Ingstad et al 1990, Wiegmann et al 2005) rather than categorizing on the basis of the phase of flight in which the crash occurred. The ‘General Flying’ phase in this study (phase three – see Figure 5.1) most closely corresponds to the ‘cruise phase’ used in previous research, although it is unclear in previous research whether ‘cruise’ includes the circuit and approach phases of flight. The finding that most fatal accidents occurred during the cruise phase (van Doorn & Zijlstra 2006) is not reproduced in this study since UK data shows that most fatal accidents occurred on ‘Launch’ (see Table 5.11) despite this being the least frequent accident phase after ‘Pre-Flight’. The discrepancy in these figures could be a result of the predominant use of aero-tow launching in the US and the popular use of winch launching methods in the UK. However the ‘General Flying’ phase in this study did contain more accidents than other categories and led to the largest number of injuries. This phase was particularly associated with the more experienced group of pilots (see Table 5.11). Previous findings that ‘Landing’ was the most frequent accident phase (e.g. van Doorn & de Voogt 2007, O’Hare 1994, AOPA 2006) were supported by the current research only if seminal events taking place in the proceeding 'approach' phase were included with those in the 'landing' phase (which may have occurred in the aforementioned study given the lack of fidelity, i.e. only three flight phases).

The significant difference in numbers of 'out-landing' accidents between the two experience groups was to be expected considering that pilots with 10 hours or fewer solo experience would rarely fly out of range of the base airfield in the UK and therefore hardly ever need to make an out-landing. This explains the finding that the rate of accidents occurring to pilots with 10 hours experience or fewer was higher than for more experienced pilots, except when the seminal event occurred during the 'General Flying' phase (Table 5.11). In any cross-country, competition or soaring flight this phase makes up the majority of the duration of the flight and therefore experienced pilots carrying out such flights are exposed to this phase for much longer. Accidents initiated in flight phase 3c ('Search/Descent') were exclusive to the more experienced group, accounting for 12% of all accidents. This was partly because most of the seminal events involved an inappropriate choice of landing area made during the 'descent and searching' phase. Landings in unplanned locations such as farm pastures, crop fields or scrubland bring dangers not associated with airfields, such as an uneven surface, obstacles and slopes, and usually a smaller area in which to land. It has previously been found that 'collisions with objects took place predominantly in terrain unsuitable for landing' (van Doorn & de Voogt 2007). These dangers are faced almost exclusively by pilots with over 10 hours experience who are more likely to undertake longer (distance) flights than are novice pilots.

All major flight phases except 'General Flying' were found to have a much higher rate of accidents for the 10 hours or fewer (inexperienced) group compared with the group of pilots with more than 10 hours as pilot in command. The highest accident rate overall was associated with inexperienced pilots and the 'Approach' phase (1 in every 2,507 launches using the mid-point estimate – see Table 5.11). This finding is not in line with general aviation, where the approach phase has been found to be one of the least likely to cause accidents to low-hours pilots (Baker et al 1996). This stark difference could be explained by the very different nature of approach control in gliders and powered aircraft. General aviation pilots use power to manage the energy on approach (speed and height). Glider pilots must set up an approach with plenty of reserve energy to reach the airfield, and then use airbrakes to “spoil the lift over a portion the wing and increase drag” (Piggott 1997), and in this way descend properly on the glide slope. In the event of an undershoot developing the airbrakes can be shut because there should be energy in reserve. However should this be mismanaged or misjudged then there is a real possibility that the glider may end up unable to reach the airfield, with no other options available. In powered aircraft this cannot happen unless there is an engine problem, so any mishandling of energy on approach can be simply rectified by the application of power or performing a go-around, a maneouvre not possible in a glider for obvious reasons. Without the ultimate option to abandon the approach (go-around) glider pilots must always correct errors occurring in this late stage of flight with little time and little height remaining. This may challenge the ability of inexperienced pilots. The causes of accidents on approach require further study.

Due in the large part to the number of accidents caused on approach, the segment of flight from joining the circuit through to the end of the approach was associated with exactly half of all accidents occurring to low hours solo glider pilots, unlike general aviation where the same section of flight was found to account for only 4% of accident initiation Baker et al (1996). The ‘Landing’ phase provided the second highest rate (after the approach) of accidents for very inexperienced pilots (1 in 3,448 launches). This supports previous findings for both gliders and aeroplanes regarding the frequency of landing accidents (O’Hare 1994, van Doorn & de Voogt 2007), as well as being in line with research findings from Baker et al (1996).

Like the approach, the landing is a discrete phase of relatively short duration, occurring late in the flight and therefore allowing little opportunity for recovery from serious errors. Landings were associated with fewer injuries than approaches which supports the finding that only 10% of landing accidents are fatal (van Doorn & de Voogt 2007).

The ‘Pre-Flight’ phase had a lower accident rate than both the ‘Approach’ and ‘Landing’ phases (one accident per 9,194 launches for the less experienced group – see

Table 5.11). This phase, however, showed the biggest discrepancy between the inexperienced and experienced pilot groups. Using the mid-point exposure estimates the less experienced pilots were over 13.5 times more likely to have an accident when the seminal event occurred during the 'Pre-Flight' phase. Pre-flight involves critical actions prior to take-off (particularly shutting the canopy and locking the airbrakes). It is possible that inexperienced pilots were actually making more of these errors, but equally possible that the difference was caused by inexperienced pilots failing to recover from errors made at this stage, since a recovery from such an error would not feature in the accident statistics.

It has been shown that accidents occurring to very inexperienced pilots (10 hours and fewer) had a very different pattern to those of more experienced pilots (over 10 hours), in terms of flight phase seminal events. Perhaps unsurprisingly, accidents occurring during training ('genuine instructional flights') followed a pattern closer to the inexperienced group than the experienced one, despite being a subset of the latter (due to the pilot in command being the instructor, and hence always with over 10 hours of experience). Instructors must allow student pilots to fly imperfectly, in order for them to practice and learn. An example of this philosophy is given in the BGA instructors' manual, which gives the following advice to instructors:

> "Unless things are going badly wrong, start with the indirect prompt. It is an invitation to the trainee to assess the situation and make a suitable decision. The prompt doesn't tell him exactly what to do" (British Gliding Association 2003)

Therefore the instructor will usually allow a situation to develop beyond what he or she would have done had they been flying solo. Hence, it is probable that had the instructor pilot been flying solo then the situation preceding the seminal events, and in many cases the seminal events themselves, would not have happened. Because it was the input from the trainee pilot that brought about these accidents, it follows that the trend would appear similar to the trend for accidents occurring when those same trainees have gone solo, and indeed this effect is apparent in the results (see Figures 5.3 and 5.4).

However this is not the full story. Within the similar patterns of accidents occurring to trainees and inexperienced solo pilots, there were some notable differences. It is evident that the approach phase led to proportionally fewer accidents during instruction (as opposed to flights by inexperienced solo pilots), and the landing and launch phases proportionally more. A probable explanation is suggested by the overall distribution of instructional accidents across the flight phases. Instructional accidents were disproportionately caused during the launch and landing stages of instructional flights compared to the other pilot group breakdowns (i.e. all pilots, 10 hours or fewer and 10 hours and over). These two phases are at the very start and end of each flight and are the only phases which include interaction with the ground. This means that, compared to the other phases, there is much less height or time for effects of seminal errors made by trainees to be corrected by instructors. Notably the approach phase was the only phase where accident initiation increased substantially for student pilots once the instructor was no longer in the aircraft (Figures 5.3, 5.4). The 'pre-flight' phase also shows this effect, but the numbers are extremely small, and hence cannot be used to suggest trends. Landing and Launch accidents declined markedly (in proportion to other types) after solo while general flying and circuit accidents remained generally similar. The unique nature of gliding undoubtedly plays some part in these effects. The instructor does not have the option to increase power in order to recover a situation such as aborting the landing attempt, or to help recover speed after release (landing and speed recovery were the two most obvious seminal event 'spikes' of instructor accidents). However the instructor would have proportionally more time to mitigate the consequences of errors made on the approach.

The high proportion of instructing accidents occurring after simulated launch failures (as a result of events during recovery after release) could be a reflection of the high number of simulated cable breaks being practiced. It could be that it is a difficult phase for instructors to recover from, once a mistake is made. It is equally possible that real launch failures are rare and so accidents caused by them are also rare. There is therefore no easy solution to this problem. It would be a possible over-simplification to use this evidence to argue for a reduction in launch failure training, in order to prevent training accidents. It could be that the amount of training itself is what leads to the low numbers of launch failure accidents when solo, and hence to take away that training would lead to a rise in solo accidents of this type. The figures could be explained by the hypothesis that because trainees are subjected to these situations many times during training the standard of solo pilots is high in this maneuver.

Conclusions

The distribution of seminal accident events among the various elements of glider flight is quite different for inexperienced glider pilots than for pilots with more experience, or for general aviation pilots. In contrast to the accidents occurring to low hours general aviation pilots, low hours glider pilots were most likely to have accidents initiated in the 'approach' phase. This may reflect a major difference between operating gliders and powered aircraft. 'Landing' also had a high accident rate for inexperienced glider pilots, in line with previous findings across many types of aviation. Approach and landing made up 68% of seminal events leading to accidents for inexperienced pilots. The group of pilots with more than 10 hours of experience also had accidents originating in these phases but at a much lower rate. In general, the approach phase accounted for far more injuries than the landing phase. All flight phases other than 'General Flying' showed higher accident rates for the pilots with fewer than 10 hours experience as pilot in command compared to more experienced pilots. Experienced pilots had a high rate of accidents originating during the 'General Flying' phase, mostly in the 'Descent/Search' sub-phase.

The pattern of accidents occurring during training is more reflective of the pattern of accidents occurring to low-hours pilots than experienced ones, with launch and landing showing up as key areas where seminal events led disproportionately to accidents, even more so than for the low-hours solo group. The approach phase however was not as prominent in instructional accidents as it was for the low-hours solo group. Therefore it seems that accidents caused by seminal events in the approach were more likely to happen once the pilot was solo and prior to them gaining experience.

This research demonstrates the benefit of examining the flight phases in which seminal accident events occur, rather than classifying accidents by the flight phase in which the crash happens. This approach allows for better targeting of appropriate remedial actions. It also shows the benefits of analyzing accidents with respect to rates rather than frequency counts or odds ratios based on accident populations.

Chapter 6 - Causal Factor Accident Analysis

Introduction

So far, it has been shown that very inexperienced glider pilots (defined as those having10 hours PIC or fewer) have the greatest risk of accident involvement (Jarvis & Harris 2007b) and that these pilots were eight times more likely than more experienced pilots to initiate accidents during the approach phase, and six times more in the landing phase (Jarvis & Harris 2008). Accidents in the approach phase were far more likely to have injuries associated with them than the landing phase. Experienced pilots were much more likely to have accidents initiated during 'general flying' activity such as ridge soaring and cross country flying, and were more likely to have accidents associated with off-airfield landings (Jarvis & Harris 2008). Eighty percent of UK gliding accidents have 'pilot related' causes (Jarvis & Harris 2008) which is line with previous studies from the US (van Doorn & de Voogt 2007).

Little research has been done on the nature of these pilot related causes in gliding accidents. Previous research carried out on aircraft accident databases containing reports for powered aircraft has used either a categorical or an analytical methodology to discover causal factors. Categorical approaches develop taxonomies where accidents are broken down into groups deemed to have circumstantial similarities. Analytical research uses taxonomies describing the underlying causal factors. Various taxonomies have been produced and utilised using both methods.

Categorical methods benefit from not imposing a pre-determined structure onto accident data, hence the methodology is predominantly data-driven. Using a categorical method the UK Civil Aviation Authority (CAA) developed ten categories of accidents including 'Controlled Flight into Terrain' (CFIT), 'Loss of Control in Visual Meteorological Conditions' and 'Low level Manoeuvring/Aerobatics', (UK Civil Aviation Authority - CAA 1997). These categories were formed from groupings of similar accidents after reviewing reports in a database. Such categories describe the general circumstances of the accidents but do not attempt to explain the human factor root causes.

These data-driven processes have spawned well-known categories which have been considered useful enough to be re-used by many organisations. An example is ‘controlled-flight-into-terrain’ (CFIT) which is often used when an aircraft is deemed to have been flying under control but hit terrain or water. But although these methods are popular, particularly within aviation safety organisations, their explanatory weakness is demonstrated by the fact that there is no agreement upon categories or definitions, even of regularly used terms; “Getting investigators and researchers to agree on what is, and more importantly, what is not CFIT, is difficult at best” (Shappell & Weigmann 2003a). Moreover, because these categorisations often give little indication about the human factor causes surrounding the accident, remedial work can be difficult.

Simple categorisation also does little to explain the human factor causes of accidents. At best, it is merely descriptive of the consequences, for example by definition CFIT is an outcome rather than a cause. In 1996, ‘The Bureau of Air Safety Investigation’ used a similar method but with more focus on human factor causes. ‘Pilot factors’ were broken down into ‘medical’ (3%), ‘improper operation of primary flight controls’ (5%), ‘inadequate pre-flight preparation or planning’ (11%), ‘perceptual misjudgement’ (12%), ‘operation beyond experience or ability’ (13%), ‘in-flight decisions or planning’ (15%), ‘did not obtain or maintain flying speed’ (15%), ‘diverted attention’ (16%) and ‘poor judgement’ (17%), (Bureau of Air Safety Investigation 1996). BASI attempted to base its investigation efforts around error frameworks such as Reason’s theory of latent failures (O’Hare 2000) and it is clearly more representative of the human factor causes than the previously described system. However it could be argued that while some of the categories chosen represent the root of human mistakes and errors (i.e. perception, decision making and judgement errors) some represent the consequences of these (e.g. ‘did not maintain flying speed’ and ‘inadequate preparation’). The latter could be caused by the former, meaning that the categorisation is questionable in terms of pilot causation, since it is used as if the categories are mutually exclusive. For example, ‘diverted attention’ could cause ‘failure to maintain flying speed’.

Recently taxonomies of accident causation based specifically on human error theory have become more popular. As early as 1982, Feggetter designed a comprehensive human factors checklist for accident investigation including cognitive, social and situational factors (Feggetter 1982). The foundation of the cognitive component was a general model of human information processing, including attention, perception, memory, decision making and response. However the checklist was very involved and intended for use by an ‘appropriately trained human factor specialist’ (Feggetter 1982). It was aimed at single accident investigations where in depth information was available, rather than multiple accident analysis based on shorter written reports, summaries or database entries.

A number of systems and taxonomies for human factor analysis of accident report databases have been developed in the last fifteen years. Gerbert and Kemmler (1986) used 1448 critical incidents reports of ‘near-accidents’ from pilots in the German Air Force to produce a taxonomy of 61 error types, fitting a four dimensional error structure of vigilance errors, perception errors, information processing errors and sensorimotor errors.

O’Hare et al (1994) coded 284 aviation accident reports deemed to have been caused by pilot error into one of three error stages based on fundamental cognitive theory; information, decision and action. It was found that 62.5% of serious and fatal accidents were characterised by decision errors. The same accident reports were coded for the existence of the 61 error types identified by Gerbert and Kemmler. A ‘principle components analysis’ found that the majority of the accidents were accounted for by seven components; ‘Mishandling controls’, ‘decisions and judgement of go-arounds’, ‘failure to check/monitor’, ‘stall/spin delayed recovery’, ‘failure to follow procedures in go-arounds’, ‘misjudgement of weather conditions’ and ‘misjudgement of altitude and clearance’.

In a second study, O’Hare et al (1994) looked at aviation accidents from 1983 until 1991, using a six step taxonomic algorithm (Figure 6.1). Of the seven human error categories formed by working through the steps, the first was ‘non-cognitive’, i.e. where the pilot had no realistic chance to intervene, such as a problem caused by a technical defect. The remaining six were information, diagnostic, goal, strategy, procedure and action error. Such a model is an expansion of the three step model previously mentioned (information, decision, action) but with the 'decision' category broken down into a further four categories. Again, the accidents with the most serious outcomes were characterised by decisions, and of the four decision categories 'goal errors' (i.e. choice of option/goal) was by far the highest (O'Hare et al 1994).

Development of a taxonomy capable of identifying deeper root causes culminated in The Human Factors Analysis and Classification System (HFACS) developed by Weigmann and Shappell (2001). HFACS is perhaps the most widely used human factors accident analysis framework, being extensively used in civil and military accident analysis in recent years (Li & Harris 2005). HFACS was developed from Reason's organizationally based model of human error (Reason 1990). The HFACS taxonomy contains four levels describing different layers of failure in the system; Unsafe Acts, Preconditions for Unsafe Acts, Unsafe Supervision, and Organizational Influences (Shappell & Weigmann 2000). Using this analytical taxonomy derived from an underlying theory, the active failures (errors) of pilots combine with latent conditions lying dormant in the system to breach its defences. These latent conditions are spawned in the upper levels of the organization and are related to management and regulatory structures.

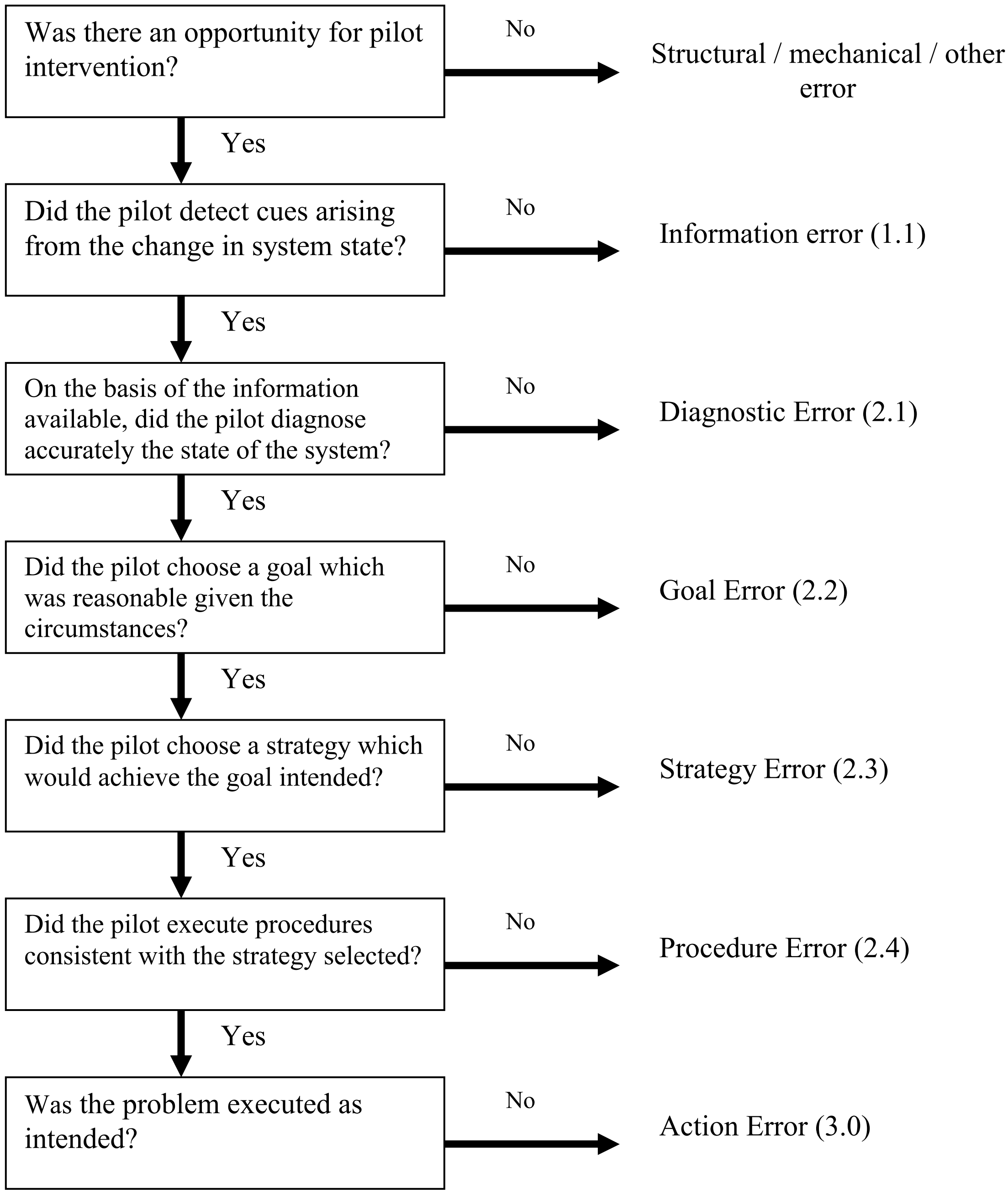


Figure 6.1. Taxonomic algorithm for classifying information processing failures (from O'Hare et al 1994)

HFACS has not been applied to gliding accidents. It has however been applied to general aviation accidents (Shappell & Wiegmann 2003a; Shappell & Wiegmann 2003b; Shappell & Wiegmann 2004). Although 19 causal categories across four levels of human failure are contained within the HFACS framework, only one of those four levels (unsafe acts of aircrew) was used to examine general aviation accident data. This level consists of four causal categories; skill-based errors, decision errors, perceptual errors and violations. Eighty-two percent of 3,200 fatal accidents were found to be attributed (at least partially) to skill-based errors; 12% included perceptual errors and 32% decision errors (Shappell & Wiegmann 2003b). Decision errors represented only about half the figure reported by O'Hare et al (1994) and the percentages of decision errors were similar for fatal and non-fatal accidents (Wiegmann et al 2005). Fine-grained analysis found that 'maintaining direction control', 'airspeed', and 'stall/spin' made up most skill-based errors (Wiegmann et al 2005) which partially supports the CAA's (1997) finding that high numbers of accidents involved loss of control. The majority of decision errors involved 'in-flight planning'; 'pre-flight planning'; 'fuel management'; 'poor selection of terrain for take-off/landing/taxi' and 'go-around decisions' (Wiegmann et al 2005).

It is arguable that none of the lower-level decision sub-categories apply to gliders in the way that they do to powered aircraft. 'Pre-flight planning' cannot be carried out in gliders to the extent to which it can be in powered aircraft simply because flights are not entirely predictable in terms of height, track and speed. Goals are often set (waypoints for example in competition flying with 'legs' drawn between them) but these are targets which the pilot hopes to achieve while accepting that s/he may not be able to, and may need to land-out (i.e. in an unknown field). However, to maximise air currents glider pilots do perform dynamic in-flight planning, described as "whole chains of logic that can play important roles in cross-country flight" (Reichmann 1992). Additionally, the circuit and landing phases require considerable dynamic planning since the vertical profile is more or less fixed, meaning the pilot must extrapolate this in order to plan to arrive at the right height, in the right place for a landing to be made. This area represents a large difference between powered aircraft and gliders because gliders cannot be flown around standard circuits with the expectation of losing a predictable amount of height all the way round (Piggott 1991). It is clear therefore that the planning issues are quite different in gliders, and therefore a read across from the HFACS categories should be treated with caution. The categories of 'fuel management' and 'go-around' are clearly not applicable to gliding operations at all. Finally, the category of poor terrain selection can only really apply to landing (rather than take off or taxiing, as in the aforementioned study). Clearly therefore the five decision categories cannot simply be mapped across and used for gliding accidents because as well as significant operational differences, the nature of decision making and judgement are very different than for pilots of powered aircraft. This supports the argument that gliders need a separate form of accident classification, and that in particular the nature of decision making and judgement need careful thought.

In addition to differences in categorisation, the upper levels of the HFACS framework are arguably unsuited to UK gliding, because it is not clear what the organisational, management or supervisory factors are for pilots operating gliders in the UK. Pilots generally fly for their own pleasure, and with very few exceptions instructors are volunteers. Recreational flights are not affected by organisational or management pressure in the way that commercial flights are. This is probably the reason that only the four lower level causal categories of HFACS (the unsafe acts of the pilots) have been applied to recreational general aviation accidents (Shappell & Wiegmann 2003a, Shappell & Wiegmann 2003b, Shappell & Wiegmann 2004). The lower level of the HFACS taxonomy relates only to the unsafe acts of the pilots, and would therefore be the only part of HFACS possibly appropriate for recreation gliding.

Despite the extent of analysis done on recreational aviation accidents, only a small amount of research has looked at very inexperienced pilots, or the differences between pilots of varying experience levels. Although no such research has been done for gliding, some has focussed on general aviation. A study of US general aviation accidents found that of 327 accidents occurring to student pilots flying solo (i.e. early solo pilots), 146 (45%) involved loss of control on landing (Baker et al 1996). The majority of these involved running off the side or the end of the runway, nosing over or hard landings. Very few led to injuries. Other prominent issues found to initiate accidents were fuel starvation (13%) and stall (13%). Stalling accidents accounted for 46% of fatal accidents identified in the study, and occurred mainly on takeoff and final approach.

The UK Civil Aviation Authority reported that accidents involving loss of control were associated with inexperience (CAA 1997). These accidents were defined as "where the pilot lost control of the aircraft while operating in accordance with normal aviation procedure, under visual flight rules (VFR), and not engaging in low level aerobatics" (CAA 1997). The mean total experience of pilots involved in LOC-VMC accidents was 927 hours, whereas the mean figure overall for all other accident types was 2115. Only 21% of pilots involved in LOC-VMC accidents had more than 1000 total hours whereas all other accident types combined had a figure of 55%. A pilot's previous amount of experience on the aircraft type involved in the accident appeared to make little difference, it was the total flying experience overall that was notable (CAA 1997). The explanation offered for this in the UK CAA publication CAP 667, states that such accidents are caused by a lack of basic handling skills often associated with an unfamiliar situation and possibly caused by initial training or inadequate refresher training (CAA 1997). It is claimed that this also brings the standard of instruction into question.

Research by O'Hare et al (1994) using the six step taxonomic algorithm (figure 6.1) found that less experienced pilots were more likely to be involved in accidents classified by the six cognitive categories (i.e. they were more likely to make human errors). This finding has not been replicated for gliding where no significant difference was found between highly inexperienced pilots and their more experienced counterparts, in terms of pilot-related causes as a proportion of all accidents (Jarvis & Harris 2008). In general aviation, the sub-category of 'procedure errors' was found to be associated with accidents involving pilots with low total hours (fewer than 100) as well as low hours on type. Procedure errors were ones where an appropriate goal and strategy had been selected by the pilot, but the procedures selected and executed were not consistent with that strategy (O'Hare et al 1994). Action errors (those where the procedures were not carried out as intended) were also more predominant in pilots with fewer than 500 total hours, and less than 100 hours on type. This is consistent with the CAA findings regarding lack of handling skills. More experienced pilots (greater than 1000 hours) suffered a greater number of information errors proportionally (O'Hare et al 1994).

In terms of human factors, a small amount of previous analysis of recreational powered aircraft accidents suggests that low hours pilots are often involved in accidents which involve general handling (skill-based) errors. There is some suggestion that instructional techniques may be to blame.

This study aims to investigate the nature of causal events in UK gliding accidents, and specifically whether any differences exist in the pattern of causal factors where those accidents involved highly inexperienced pilots. It has been argued that the human factors taxonomies developed for powered aircraft accident analysis are not completely suitable for analysing gliding accidents because of the unique nature of glider operations. In particular, pilot decisions are of a very different nature to those made by powered aircraft pilots. Therefore the analysis will be approached from a fresh angle.

Method

Before any causal analysis could be done on the accident database, a taxonomy had to be produced with which to do so. Therefore the first objective was to create a taxonomy specifically for categorising pilot related causal accident events in gliding. This was done using qualitative means to analyse the meaning of all accident events identified in Jarvis and Harris (2008), see chapter five. The second objective was to code all seminal events by the taxonomy and apply quantitative analysis in order to discover where the particular dangers are for glider pilots, and look specifically at the problems being faced by very inexperienced pilots. These objectives will be termed stage one and stage two.

Stage 1 - Creation of the taxonomy

As argued, gliding accidents may be quite different in nature to those of other aircraft and so a categorical data driven approach was chosen for the analysis of accident narratives, rather than relying on any predetermined taxonomies. This was in order to avoid the risk of overlooking elements unique to gliding. The analysis borrowed from principles of grounded theory methodology (Glaser & Strauss 1967) but also had to be capable of categorising accident events in a mutually exclusive manner in order to facilitate quantitative analysis.

Grounded theory was developed by Barney Glaser and Anselm Strauss in 1967 and is a method of deriving theory inductively from systematic analysis (Strauss & Corbin 1990). The method allows meaning to emerge from data, by constant comparison between individual sentences or phrases within the data in order to assess their meaning within the context, before grouping (coding) those phrases on several levels. Hence it is a technique that attempts not to impose any pre-conceptions onto the data. Using the principles of grounded theory allowed a structured approach to be taken to the qualitative data analysis.

The approach taken followed three stages; Data gathering, open coding and axial coding, as shown in Figure 6.2.

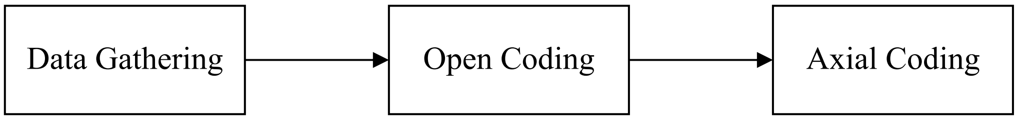


Figure 6.2. Initial grounded theory application steps (From Huddlestone and Harris 2006)

Data Gathering

Only UK gliding accident reports from a five year period (2002 to 2006) identified by Jarvis and Harris (2008) as having a pilot-related cause were used in the analysis (source: British Gliding Association 2007b). Each accident report contained details such as glider type, pilot age and experience; injury, damage caused, and a narrative description of what happened. Within the 359 reports, 545 contributing events had previously been identified Jarvis and Harris (2008).

Where possible the narrative accident descriptions contained within the BGA database were supplemented with segments from AAIB or BGA reports, as well as additional information from other fields in the database (with the strict exception of pilot experience for the purposes of coding). In a very few cases witnesses or pilots also contributed information about specific accidents. Hence for each accident a 'descriptor' was created that, in a number of cases, contained more information than the original BGA database text description.

Open coding

Strauss and Corbin (1990) describe open coding as "The process of breaking down, examining, comparing, conceptualising and categorizing data". In terms of 'concepts', these are defined as "conceptual labels placed on discrete happenings, events and other instances of phenomenon" (Strauss & Corbin 1990). Since any individual accident is naturally a 'discrete happening', there was no need for the data to be broken down in this way, as occurs commonly in grounded theory analysis. The only breakdown was splitting accidents into events which were also discrete happenings. This was done in

Chapter five; accident reports were broken into discrete events deemed to be pilot-related. With the data already reduced ('broken down'), a modified version of open coding was performed on all accident events used using pair-comparisons, as described by Partington (2002). Each data item (event) was compared with preceding data items to see if it described the same phenomenon. Where this was deemed to be the case they were coded as belonging to the same category.

For example, the first event to be looked at was a description of an accident where a pilot forgot to lock the canopy and it came off during the launch. The next was an event where the pilot misjudged the approach path and allowed a serious undershoot to develop. Comparing the two, it was decided that they were dissimilar in terms of their underlying causes and hence two coding categories were started but no definitions or labels were applied at this point. The third event was a pilot not doing their straps up prior to launch. This was compared with both the previous events and declared to be similar to the first one, and so put into that pile. A new category was started whenever an event was deemed to be dissimilar to all preceding events up to that point. The process was repeated for all 359 accidents and the categories were continually revised and re-worked. Only after the process was complete was each category given a working title (code).

The objective of the research was to compare accident events with each other, and so each accident event necessarily remained separate as a phenomenon. In other words rather than pulling out as many themes as possible from every accident and then categorising these themes, the categorisation process was performed on individual accident events, each given a single overall theme. Importantly, by rigidly applying only a single code to each accident event, quantitative analysis of events/accidents remained a valid procedure, and so comparison of accidents was possible after the qualitative analysis was complete. This was because every accident event had the same number of associated themes (one), and therefore there was no numerical bias caused by the applying the qualitative methodology.

Open coding was performed on accident events using constant comparisons between all 545 event descriptors (treated as 'concepts' in terms of grounded theory principles). As previously stated the descriptors were compared very strictly on a single criteria: the basis of the pilot contribution to the accident. Other criterion could have been applied, for example: outcome, consequence, 'type' of accident, damage caused etc, and in many cases these criterion would have been a more obvious choice. In other words each descriptor was compared only in terms of what the pilot did or did not do. In this way, confusion about cause and consequence was avoided. Furthermore the categories were constantly and systematically compared to make sure each was mutually exclusive, in order to reaffirm that a single criteria was being applied. Where there was found to be overlap between categories, they were re-worked with particular care given to the criterion being applied.

After all 545 events had been grouped using this open coding process, each of the groups (categories) were labelled on the basis of its descriptions represented in terms of pilot actions/inactions/activities. This process involved a second subject matter expert, in an attempt to reduce any potential researcher bias. Additionally some events were re-categorised at this point, after further discussion. Sixty-five categories emerged from this open coding process.

Axial-coding

Axial coding is defined as "A set of procedures whereby data are put back together in new ways after open coding, by making connections between categories", Strauss & Corbin (1990, p.96). In pure grounded theory analysis a process called paradigm linking is used, which links categories by causes, actions and consequences. However since all events had already been carefully coded on the basis of cause, this was inappropriate. Comparison of categories was used however, in order to link together those categories that had similar human factor root causes. For example many categories had attentional issues as the root cause of the events (e.g. action omitted; not lowering gear or flap, action slip or not noticing another aircraft, etc). These categories were then linked to form a major human factors category. Additionally some categories were split further to reflect different root causes of a similar issue. For example, one of the initial categories was called ‘flew glider into poor weather/ gliding conditions’. It was clear from scrutiny of the reports that in some cases this had been an attentional issue, for example where a pilot had not noticed a change in wind direction, while with some it had been a strategic decision, for example where the pilot flew into rain and low visibility in an effort to find lift. Therefore these were separated into two categories to reflect the difference in cause. Using these methods, a hierarchical template was arrived at that described all gliding accidents within the database.

In this way the categories from the open coding process were systematically compared and formed into larger groups, based on pilot action/inaction. Some of these groups were then joined to form larger units, and finally six higher level categories emerged describing the root human factor causes. Hence the taxonomy ended up with a hierarchical structure. After all the further coding and grouping, the number of categories decreased to 62, and these became the lower level categories.

Selective coding, sometimes used as a final process in grounded theory analysis in order to describe the central themes of the ‘story’ was unnecessary, since when the categories and sub-categories were arrived at, numerical and statistical analysis was to be applied. This numerical analysis was valid because of the rigorous application of single codes to single events. Since total numbers of events were known for each category and sub-category, and each event was only coded once, quantitative analysis could be performed on the accident database using the hierarchical template.

Because total numbers of events were known for each category and sub-category, and each event was only coded once, quantitative analysis could be performed on the accident causes.

Reliability and blind rating

A number of measures were put in place to insure the reliability and objectivity of the rating process and provide triangulation of the categorisation, in common with the previous study (see chapter 5).

- First, pilot experience information was removed before rating the reports to avoid any possible bias during the grounded theory analysis.
- Second, the complete categorisation along with exemplar narrative statements were given to a number of subject matter experts for feedback as to whether or not they appeared valid and reflective of the activity, and some adjustments made in light of this exercise.
- Third, after coding was complete one hundred complete accident narratives were randomly selected to be used as the reliability sample (representing over 25% of the narratives and containing 150 events in total). Each event was highlighted within the report in order to ensure that the second rater was making a judgement based on the correct event. Human factors training was provided to the second rater, who was a subject matter expert in gliding, a former airline training captain and a qualified instructor in crew resource management. The template was explained thoroughly and practices were performed. The independent rater then proceeded to categorise all 150 events using the template. The complete inter-rater grid is shown in Appendix C part 1, with the two sets of 150 codes (for rater A and rater B) in Appendix C part 2.
- Fourth, an intra- rater reliability test was conducted in the same way on the same sample (as the inter-rater test) two weeks after initial rating was completed. Intra-rater reliability (observer consistency) is used to establish that the researcher's method of coding is consistent, and inter- rater tests are used to make sure that the researcher had not coded in an idiosyncratic way (Robson 2002). The complete intra-rater grid is shown in Appendix D part 1, with the two sets of 150 codes (for 1st and 2nd rating) in Appendix D part 2.
- The lower level categories were matched to the four top level categories by the independent rater, and the vertical structure of the template was discussed in

light of this exercise. Some modifications were made to terminology of several lower level categories as a result.

Statistical analysis

Chi square and Fishers exact tests was used to compare accident features between groups. These test the assumption that certain properties of groups are randomly distributed between those groups. In the case of accident analysis (particularly analysis of cause rather than outcome), accidents can be treated as a random sample of all flights undertaken, until evidence becomes available to indicate that certain features of accident flights were different to other flights. In other words, the population is all flights undertaken, the sample is all accidents. Fisher's exact tests can provide evidence of imbalance between accident features belonging to the different pilot groups. Thus if a p value is significant it would suggest that the variables under consideration are systematically associated with accident causation, within the population of all flights. Rates for particular accident factors were calculated in the same way as used by Jarvis and Harris (2008), using exposure data from the years 2002 to 2006 inclusive.

Results

Reliability

The proportion of agreement for the intra-rater test on the low level categories was 0.913, giving a Cohen's Kappa of 0.909. The inter-rater test gave a proportion of agreement of 0.773, with a Cohen's Kappa of 0.764. These reliability coefficients can be regarded as 'excellent' (>0.75) according to Robson (2002). When this same test result is analysed between just the six high level categories, the figures increase only slightly, to Kappa = 0.94 (intra-rater test) and Kappa = 0.78 (inter-rater test). The similar coefficients show that there is only a small amount of inconsistency in selection of the finer-grained elements (lowest level categories) within the larger six categories (i.e. that most of the inconsistency is *between* the higher level categories).

The coding of the template can therefore be considered as reliable and consistent. However the relation of specific lower level categories to top level (general) categories required verification. All 62 lower level category descriptions, along with an event example were matched to the four main general categories by an independent rater. No intra-rater test was conducted because the investigator was familiar with the template. The results were a Kappa of .64 for the inter-rater test (Good). It was of interest to note that the majority of disagreement was between perceptual judgement (J) and handling (H). This could have been due to the descriptions given to the second rater.

Table 6.1 (across two pages). The final template with accident numbers for both groups, and injuries. Each accident is represented by its seminal event only. The table is presented in four sections, representing the four top-level groups. Where a number of categories from open coding were brought together into intermediate categories, those are shown on the left in grey.

	J Perceptual Judgement	=< 10h	>10h	Injuries
Misjudged Vertical Separation	Misjudged intended separation with obstruction	0	2	1
	Allowed wing tip to touch the ground in a turn	0	2	0
	Landing Flare too high/ too early	4	8	0
	Landing Flare too low/ too late	3	28	2
	Misjudged lateral separation from object	0	3	1
Misjudged positioning in circuit	Flew too close in / (too high / too much energy)	1	11	2
	Flew too far out/ (too low / little energy)	1	10	1
	Misjudged alignment of final turn exit	0	0	0
Approach Misjudgement: Undershoot	Opened airbrakes at inadvisable point in approach	0	2	0
	Left airbrakes out too long	1	3	0
	Insufficient reduction of airbrakes below glide path	3	2	1
	Applied landing flap too early	0	0	0
	Not known	1	3	1
Approach Misjudgement: Overshoot	Used too little airbrake	0	0	0
	Put airbrake away (any number of times)	0	1	0
	Continued until too late to safely execute alternative	0	1	0
	Cue misinterpreted	1	1	0
	H - Handling	=< 10h	>10h	Injuries
Mishandled in Pitch	Mishandled elevator causing ‘longitudinal oscillation’	0	4	1
	Overuse of up-elevator causing accelerated stall	0	3	2
	Overuse of up-elevator causing overly steep climb	0	2	1
	Overuse of up elevator causing tail strike in flare	0	3	0
	Overuse of down elevator	0	7	2
Mishandled in roll	Not correcting/allowing for drift	2	1	1
	Mishandled sideslip	0	2	0
	Mishandled airbrakes	1	3	2
	Mishandled flaps	0	2	0
	Mishandled during an established aerotow	0	3	1
Ground run after landing	Allowed/ failed to prevent wing going down	0	16	5
	Steering error (NOT misjudged lateral proximity)	1	4	0
Mishandled Pitch during ground run	Causing a re-light after landing by mishandling controls	0	1	1
	Causing a premature/exaggerated take-off on launch	0	2	0
	Causing an extended ground run on take off	1	2	0
	Overuse of wheel brake	0	0	0
Mishandled speed	Failed to maintain (or increase to) speed required	3	25	10
	Allowed unintentional increase of speed	1	3	0
	Changed hands on controls	0	1	0

S - Strategy		=< 10h	>10h	Injuries
Continuing with a plan / strategy	Continued into poor conditions	0	4	1
	Accepted launch into unfavourable conditions	0	3	1
	Continued with a marginal attempt to reach airfield	0	7	2
	Continued operating (e.g. thermal soaring) while too low	0	8	2
	Flew over an area with no safe landable options	0	7	1
	Continued with compromised launch	0	3	0
Flew out of reach	Flew out of reach of airfield during intended local flight	0	20	7
	Flew out of reach of chosen field while x-country/ soaring	0	6	1
	Compromised sighting of intended flight path	0	2	1
	Landed in an unsuitable field	0	12	1
	Chose unsuitable initial field leading to late rejection	0	6	1
	Chose to land on unusual/unfamiliar area of airfield	0	5	0
	Rejected straight ahead landing from launch failure	0	1	1
	Deliberately left wheel up	0	1	0
	Direction of landing (into land-out field) unsuitable	0	1	0
A - Attention		=< 10h	>10h	Injuries
Did not notice stimulus	Did not notice obstruction/ditch/undulation/slope	0	8	0
	Did not notice another aircraft, or launch in progress	0	8	4
	Did not notice change in conditions	1	5	1
	Aircraft issue/setting overlooked	1	3	1
Secondary control actions	Secondary control action omitted (unintentionally)	0	29	1
	Action initiated but not completed. E.g. u/c, canopy lock	0	2	1
	Secondary control action slip	0	16	2
	Failed to correctly set/lock control/straps/seat	2	9	3

Four accidents were dropped from the analysis because either the cause could not be sufficiently determined, or the accident was caused by a non-handling passenger. These represented three open coded categories in their entirety and so these were dropped from further analysis, leaving only 59 low level categories in the taxonomy. All 59 are listed in Table 6.1 along with the main intermediate subgroups formed subsequently (on the left). The template is in four sections, each representing one of the four top-level groups that finally emerged from the axial coding process: Perceptual Judgement (J), Handling (H), Strategy (S) and Attention (A).

The four top-level groups were formed as follows.

Perceptual Judgement (J): Categories in this group included those with events driven by perceptual judgment that turned out to be in error. For example misjudged separation (e.g. with object), misjudged position, rate, descent profile or height. Perceptual judgement was not categorised alongside decision judgement (e.g. the judgement of alternatives).

Handling (H): Categories forming this group included events deemed to be caused by mishandling of the controls (e.g. under-banked turns, over-pitching the launch rotation, overuse of wheel brake, etc.). This is where the outcome of control actions was not as intended, in contrast to 'perceptual judgement' or 'attention' events.

Strategy(S): This included all eventualities where the glider's overall operation led to the accident in a manner fitting the description of 'pilot-related' and included any deliberate action/inaction (as opposed to outcome) not caused by a perceptual misjudgement or handling of controls.

Attention (A): This group was formed of lower level categories where it was clear from the accident narrative that the pilot had been unaware of his/her actions (or inactions) leading to the accident, for example 'secondary control action slip' (such as operating the flap lever rather than the undercarriage lever). This category was also used where a problem occurred due to attentional issues relating to other information, such as not noticing a change in conditions (such as wind direction).

Results Stage 2 – Quantitative analysis of accidents.

For quantitative analysis, each accident was represented by its seminal event only. Being the first event in the accident sequence, the seminal event represented the cause of the accident. All analysis in the following section is based on seminal events only.

Injury causes

Table 6.2 shows the distribution of accidents and injuries between the four top-level groups. Chi-square analysis on this data (injury verses no injury) resulted in a chi square of 11.193 ($p < 0.05$, df = 3). This indicates that the 'perceptual judgement' group (J) contained significantly fewer injuries than expected whereas 'handling' (H) contained more.

Table 6.2. Breakdown of the top-levels of the taxonomy by injury and non-injury.

Top level groups	Accidents causing injury	Accidents not causing injury	Total
Perceptual Judgement (J)	9	83	92
Handling (H)	26	67	93
Strategy (S)	19	67	86
Attention (A)	13	71	84

The two most common low level categories (leaving the landing flare too late / too low, and omitting a secondary control action) were unlikely to be associated with injuries (Table 6.1). The third most common open code (not maintaining or increasing airspeed) was much more likely to be associated with injury (over a third of such accidents). This category contained more injuries than any other, including fatal and serious injuries. Flying out of reach of the airfield during a local flight was associated with the highest number of accidents leading to injuries categorised as serious or fatal. This category was applied where it could be deemed that the pilot intended to fly 'locally' (within range of the home airfield) but got into a position whereby a return was able to be achieved.

Quantitative analysis of accidents by experience

Seminal accident events in the four top-level groups were compared between the two pilot experience levels using chi-square analysis supported by standardised residual scores, and rates calculations. Where differences were found, the lower level categories were examined in an attempt to identify more detailed explanations.

Table 6.3. Seminal accident event numbers and rates by experience level for the four top level groups. The number of accidents is given for each group (n), rates are given in the form of launches per event, and the 95% confidence interval for the inexperience pilot level (estimate) is also given (+/-).

Seminal event totals and rates	10 Hours and under			Over 10 Hours	
	n	Launch per event	+/-	n	Launch per event
Perceptual Judgement (J)	15	1995	358	77	20,907
Handling (H)	9	3325	597	84	19,164
Strategy (S)	0	n/a	n/a	86	18,719
Attention (A)	4	7481	1344	80	20,123

Rates for particular accident factors were calculated using the revised exposure measure from Jarvis and Harris (2008) which was an estimate of the number of launches and hours flown by UK glider pilots between 2002 and 2006. The '10 hours and under' group totals were subtracted from the overall UK totals to provide data for two mutually exclusive groups; pilots with 10 hours or fewer PIC and those with more than 10 hours PIC. The estimated number of launches taken from 2002 – 2006 by pilots with 10 hours PIC or fewer was 29,924 (upper 95% confidence interval boundary = 35,301; lower 95% boundary = 24,548). The estimated number of hours flown was 11,553 (upper 95% confidence interval boundary = 14,017; lower 95% boundary = 9,089). The mean calculation of flying exposure by pilots with over 10 hours PIC during the same period was 1,609,810 launches and 696,041 hours.

Comparison of accident totals (the data in the two columns headed 'n', Table 6.3) gave a Chi-Square of 17.875 (df = 3, $p < 0.001$). No expected cell value was less than five and so the assumed criteria for Pearson's chi square test was upheld. Standardised residuals showed that for the 10-and-under pilot group there were significantly more seminal events associated with perceptual judgement ($z = 2.9$) and significantly fewer associated with strategy ($z = -2.6$) than for the over10 group. Accident rates support this general picture, but also show the real extent to which inexperienced pilots were much more likely to have an accident associated with seminal events in all top-level groups apart from 'strategy'. The lowest rate for seminal events associated with perceptual judgement in the under10 pilot group (1995 + 358) occurred at 8.9 times the rate per launch than for the over10 group, although none of these led to injury. Seminal events in the 'handling' category occurred at a rate 4.9 times higher for the under10 group (lowest estimate) and those associated with 'attention' at a rate 2.3 times higher. No seminal events involving strategy occurred to any pilots with ten hours of experience or fewer.

Out of the fifteen 'judgement' related events that occurred to pilots with10 hours or fewer, three were associated with using too much airbrake (insufficient reduction of airbrake) below the glide-path, four were associated with flaring the glider too early and three flaring too late. Hence nearly half of the judgement errors were misjudgements of the landing flare. The remainder were spread between a number of low level categories. Caution should be used however since the number of accidents is very low. Further analysis was undertaken on the low-level categories that accounted for the largest numbers of accidents for inexperienced pilots, three of which were from the judgement group (J). Cross-tabulations were prepared in order to compare the numbers of accidents occurring to inexperienced and experienced pilots in each of these low level categories. Fisher's exact tests were used along with calculations of odds ratios.

Table 6.4. The four most frequent lower-level categories for inexperienced pilots, broken down for comparison with the experienced pilot group. The four 2 x 2 cross-tabulations are shown by the columns '10 and under' and 'over 10'. Each totals 359 (the number of accidents overall).

	10 and under	Over 10	Fisher's exact	Odds Ratio	95% confidence interval odds ratios
Landing flare too high / early	4	8	p=.0095	6.73	1.89 - 23.96
All other accidents	24	323			
Landing flare too late / low	3	28	p=.722	1.3	0.37 - 4.57
All other accidents	25	303			
Insufficient reduction of airbrake below glide-path	3	2	p=.0038	19.74	3.15 - 123.7
All other accidents	25	329			
Too little airspeed	3	25	p=.47	1.47	0.41 - 5.2
All other accidents	25	306			

Table 6.4 shows a statistically significant difference between experienced and inexperienced pilots in terms of the numbers of accidents caused by 'insufficient reduction of airbrake below the glide-path' and 'landing flare too high/ too early'. The odds of an accident caused by an insufficient reduction of airbrake were nearly 20 times greater for pilots with 10 hours or fewer (over three times greater using the lowest 95% confidence interval). Although statistically significant, these numbers are small and should therefore be treated with some caution. However it is of note that there is only one case associated with the insufficient airbrake reduction occurring to a pilot with more than 50 hours total experience in the whole accident set. This supports the assertion that these types of events were highly associated with pilot inexperience. Accidents involving 'landing flare too high' were over six times more likely to occur to inexperienced pilots. However none of these caused injuries.

Applying the flight phase codes of Jarvis and Harris (2008) to the present categorisation showed that all but one 'landing' accident was judgement related. Of the 11 'approach' accidents identified as occurring to low-hours pilots by Jarvis and Harris (2008), the present categorisation showed that six were caused by handling issues (H), and five by judgement problems (J). All these were accounted for by just three intermediate categories. These were 'Failing to maintain (or increase to) speed required' and 'mishandled in pitch' (within 'H'), and 'Approach misjudgement: undershoot' (within 'J'). The latter was used where the approach path was misjudged to the point where the glider failed to reach the intended landing area, meaning that the airbrakes were not reduced in order to check the rate of descent. Three of these were coded as 'insufficient reduction of airbrake below the glide path (Tables 6.1 & 6.3), one as 'left airbrakes out too long' and one as 'reason not known'.

In summary, low-hours pilots were far more likely than more experienced pilots to have accidents caused by problems stemming from judgement and handling issues. Insufficient reduction of airbrake on approach was highly associated with inexperienced pilots, as was misjudgement in the landing flare.

Instructional accidents.

Jarvis and Harris (2008) identified 34 accidents that occurred during instruction. Analysis of these using the causal template found that only two fine grained codes regularly appeared as seminal events. These were 'landing flare too late to low' and 'overuse of down elevator', both of which were found to be seminal events in four accidents. Interestingly, no instructional accidents occurred due to landing flare being initiated too early, or because of overuse of airbrake on approach. This is noteworthy given that these errors caused a number of accidents to early solo pilots over the period 2002 - 2006.

Discussion

The taxonomy

The methodology resulted in a multi-layered taxonomy similar to that of Gerbert and Kemmler (1986) with a striking similarity in the number of low level categories as well as top-level groups. However the four top-level human factor groups were more in line with the basic human information processing model, paralleling the cognitive component of Feggetter's (1982) model and O'Hare et al's (1994) classification system, despite the method used to arrive at the categorisation being quite different. Unlike these taxonomies, the main groups were not constructed from cognitive theory, but were formed by classifying technical descriptions of pilot related events from accident narratives, in a similar way to Gerbert and Kemmler's (1986) research. For example, although the 'strategy' category contains mainly decision-related events, it does not break down decisions further into diagnostic, goal, strategy and procedure in line with O'Hare et al's (1994) decision error category. Instead it breaks into many smaller categories describing events from an operational perspective such as 'flying out of reach/position/sight of intended landing area' and 'accepting launch into unfavourable conditions'. This is more in line with Gerbert and Kemmler's work in which low level categories were also characterised by descriptions of pilot actions / inactions such as 'inadequate pre-flight preparation' and 'poor throttle control'. The fact that the top-level groups still aligned with models such as those by O'Hare et al (1994) supports the validity of applying taxonomies based on the human information processing system to accident narratives.

The categorisation of top level groups also has some parallels with the 'unsafe acts' level of the HFACS taxonomy, but there are some notable differences. In HFACS, 'perceptual errors' are those that "occur when sensory input is degraded or unusual, as is the case with visual illusions and spatial disorientation or when aircrew simply misjudge the aircraft's altitude, attitude, or airspeed" (Shappell & Weigmann 2000). Three quarters of general aviation accidents in the HFACS 'perceptual category' are made up of misjudged distance, flare, misperceiving altitude, misjudged clearance and visual / aural perception (Shappel & Weigmann 2004). Such events were classified under the 'Perceptual Judgement (J)' group of the present study. Additionally however, group J was made up from a number of lower level categories formed from events where pilots had misjudged more complex aspects uniquely related to the gliders. For example; a pilot flying a circuit with insufficient energy to recover to the airfield (combination of height, distance and possibly speed), or a pilot failing to recognise a developing undershoot until too late to recover. These elements of judgement are not fully accounted for by the HFACS taxonomy, even in its fine grained form for general aviation, validating the need for the bespoke approach to gliding accidents. This could also account for the reason why a substantially higher proportion of events were found to be related to 'perceptual judgement' in the present research, relative to the low proportion found in general aviation (Shappell & Wiegmann 2003b).

Decision errors as classified by HFACS are broadly equivalent to 'strategy' related events in the present study. Decision errors are said to represent intentional behaviour that proceeds as planned, yet the plan itself proves inadequate for the situation (Weigmann & Shappel 2003). The limitation with the 'strategy' category in the present study is that it was used for a number of accidents which, although undoubtedly a result of the pilot's intentional strategy, may have had other contributory factors (such as attentional and perceptual contributions, and also an element of chance). An example would be a glider losing height in an area of poor fields. The term 'strategy' was therefore applied so as to cover all eventualities where the glider's overall operation initiated the accident in a manner fitting the description of 'pilot-related' as previously given. This included occasions where it was not possible to know what (if any) decisions had been taken by the pilot, but it could be discerned that the glider was flown into a dangerous situation, albeit that the risk may have been underestimated the pilot.

The other two categories from the present study ('attention' and 'handling') have no direct parallels in HFACS but do share some characteristics with the 'skill based error' group. Skill-based behaviour is that which occurs 'without significant conscious thought' (Shappell & Weigmann 2000) and in pilot related events such errors would usually be associated with control inputs ('stick and rudder skills', Shappell & Weigmann 2000). However such categorisation may not be adequate to describe all events caused by the incorrect use of controls. For example, a glider pilot who recognises an undershoot but raises the nose rather than closing the airbrakes (hence leading to low airspeed) may have consciously taken that action, i.e. the control input was a ‘mistake’ rather than a ‘skill-based’ error. By O’Hare *et* al’s (1994) algorithm such an error would be deemed a ‘strategy’ error, since the change of state was correctly identified but the strategy to address it failed. However, one might argue that the pilot’s strategy was to extend the glide but the control input did not achieve the strategy chosen. It was not always possible to determine exactly what the pilots’ strategies and plans were from the accident narratives alone. However, where sufficient detail existed it was possible to determine whether or not the problem stemmed from the pilot mishandling the control inputs. Many low level codes were applied to events where there was evidence that incorrect control inputs initiated the accidents (e.g. ‘overuse of up elevator causing tail strike in flare’). Where this was the case, the codes were grouped and the resulting top level category of ‘handling’ emerged. Importantly, this did not represent an attempt to judge the adequacy of the cognitive process involved but simply represented a collection of events where the pilot used incorrect control inputs to solve the problem.

Low level codes involving attentional deficit, distraction or where a skill-based error was indisputable, were brought together into the top-level category of ‘attention’. For example a pilot who continually retracted and extended the undercarriage on approach with no airbrake usage could confidently be deemed to be experiencing a skill-based error, since the only reasonable explanation would be that the pilot had mistaken the undercarriage lever for the airbrake lever. Also, narratives where pilots forget to lower the undercarriage were classified under ‘attention’ since these were deemed to be lapses.

No accident events were classified as violations from the qualitative analysis. All were found to be better placed in other categories.

It was previously argued that analysis of gliding accidents would benefit from a bespoke taxonomy, involving the use of emergent methods, rather than relying on existing categorisations from powered aviation. It has been shown that the four top level 'human factor' groups that emerged from this process were in fact broadly similar to some previous classifications. However many of the underlying categories from which they were formed had no parallel in powered aviation or previous research, and some of these were very prominent in terms of accident totals and severity. Notable among them were 'flying out of reach of the landing area on a local flight', accounting for the highest number of serious and fatal injuries of any single low-level category. 'Allowing the wing to go down' accounted for a high number of 'handling' accidents on both landing and launch This issue is almost unique to gliders due to the central single-wheel undercarriage design necessitated by the long slender wings. Furthermore six categories specifically concerned the pilots' use of airbrakes. Because gliders do not have their own power, a glider pilot uses airbrakes to control the rate of descent on approach (Stewart 1994) unlike powered aircraft in which combinations of power and attitude are used (Thom 1997). These findings vindicate the lengthy process of forming a new taxonomy, specific to gliders.

Analysis of Causal Events Overall

Handling related events were more likely to lead to injury than other types, unlike previous research in which 'decision making' was found to be the most associated with injury causing accidents (O'Hare et al 1994). However it is possible that the origin and nature of the 'handling' category had some effect on this, since it may well have contained some accidents that would have been placed in the decision category of 'strategy' using O'Hare et al's algorithm (O'Hare et al 1994). In such a case, poor handling could be deemed to constitute a reason for not achieving an appropriate goal in answer to the question "Did the pilot choose a strategy which would achieve the goal intended?" (see Figure 6.1, step five), and hence this would spawn an error labelled 'strategy'. In the current study, any error involving mishandling of the controls was placed in the 'handling' (H) category. Perceptual judgement (J) accounted for a significantly lower number of injuries than the other three groups. This is partly explained by the fact that over a third of the events were misjudgements of the landing flare; events unlikely to be serious, with only two out of 31 such accidents ending in an injury. However the overall proportion of events related to 'perceptual judgement' was higher than found in general aviation (Shappell & Wiegmann 2003b). This probably relates to the manner in which gliders are flown; with pilots having to rely heavily on the judgement of visual references rather than being able to set specific parameters such as altitude, power and position.

Accidents caused by handling events involving low airspeed were not only common but also relatively likely to cause injury. Indeed 15% of all injury causing accidents had a seminal event coded under this single low-level category. This supports findings from general aviation where 'airspeed' accounts for 19% of fatal accidents (Weigmann et al 2005).

Analysis of Causal Events by pilot experience

Very inexperienced pilots suffered more accidents in all top level groups apart from 'strategy' (S). The full extent can be seen by the rates calculations. This is probably a reflection of the fact that pilots with 10 hours or fewer do not have the experience or qualifications to be permitted to fly strategically demanding flights such as cross-country tasks or competitions, and hence do not have the same opportunity to engage in difficult strategic decision making. Most very inexperienced pilots must fly within range of their home airfield, and in moderate conditions, until they have qualifications such as the British Gliding Association Bronze Badge (Jarvis & Harris 2007b). By way of illustration, the lower level category of 'chose an unsuitable land-out field' only contained accidents involving experienced pilots, probably because field selection is a task usually associated with cross country flying.

Rates show that inexperienced pilots had nearly nine times as many accidents involving judgement in the seminal event than experienced pilots. The high number of landing flare misjudgements was a major contributor, but none of these accidents were serious. Insufficient reduction of airbrakes on the approach was an initiating accident event almost exclusive to inexperienced pilots. This category was assigned to a situation where the pilot failed to appreciate that the glider was undershooting and continued to hold the airbrake lever back. This phenomenon in early hours pilots is consistent with research by Kasarskis et al (2001) who used a flight simulator (of a powered aircraft) to show that novice pilots tended to land short of the optimal landing point compared to experienced pilots. It is therefore likely that inexperienced pilots are either slower to judge when the aircraft is undershooting the desired path or that experts are better able to take appropriate corrective action. Baker et al (1996) found that only a very small proportion of accidents occurring to solo student pilots of light aircraft were initiated in the approach phase; eight out of 327 accidents (compared with 99 on landing) which is not consistent with the findings of the present study. This could be because a poorly judged approach in a powered aircraft (e.g. undershooting) can be easily escaped from by application of power, or even a go-around, and hence such events are less likely to produce accidents. In gliding, seemingly harmless errors of this sort can become very serious because they can quickly develop to the point where the landing area cannot be reached, leaving the pilot with no other options available. The skill of judging the approach path is therefore arguably a more critical task for a student glider pilot than a student in a powered aircraft, and may require more training and practice. Indeed Piggott claims that the ‘approach is a busy time for the beginner and it is only some time after soloing that he learns to organise the thinking and flying so that there is enough time for refinements’ (Piggott 1997).

Instructional accidents

Leaving the landing flare too late accounted for a number of accidents during instruction. This is unsurprising for two reasons. Firstly because early solo pilots were shown to suffer from this particular error and secondly because it is an error of omission that gives an instructor very little time to react, since it occurs close to the ground and moments prior to impact. It is clear from the analysis that a number of pilots do not fully master this skill and suffer these types of accidents while solo, but that as experience is gained the rate of accidents caused by flaring too late decreases substantially. No instructional accidents occurred due to overuse of airbrakes on approach or flaring too early, despite these being associated with accidents to early solo pilots. This could be because these types of errors give instructors time to take over, or provide verbal prompts. It is therefore possible that such errors occur as frequently on instructional flights as they do on early solo flights, but do not lead to accidents.

Conclusion

Based on the literature, it was argued that because gliding was a unique aviation activity, emergent methods based on grounded theory techniques were required to create a new taxonomy, rather rely on existing ones for powered aviation. This approach was validated by the large number of codes that emerged from open coding that had no parallel with previous research. Issues unique to gliding emerged, such as the need to cater for complex perceptual judgement events involving glide performance, and differences in accident type and number initiated during the approach phase. The top-level categories formed from these codes were broadly in line with previous human factors taxonomies, supporting the validity of the top level categories developed.

Using the taxonomy, it was found that inexperienced pilots were particularly vulnerable to accidents associated with misjudgements, chiefly in the landing flare and the approach undershoot, as shown in Table 6.4. In general however, accidents caused by misjudgement were less likely to cause injuries than those involving the mishandling of controls, primarily because landing flare accidents were rarely severe. Approach control (including misjudgement of approach path, misuse of the airbrakes and mishandling of airspeed) was found to be the most common danger for low-hours pilots. Instructors therefore need to pay more attention to this area when deciding that a student pilot is competent for solo flying.

This research demonstrates the benefits of producing accident rates rather than relying on simple frequency counts of accidents. Although statistics such as Chi square can be used to compare existing accident totals between pilot groups, the likelihood of those accidents occurring can only be demonstrated using exposure data to produce accident rates. Odds ratios and cross tabulations can give probabilities and risk of accident types occurring only on a relative basis within a given set of accidents. However, rates based on exposure data (all flying done) can show the comparative chances of accidents occurring within the population of all flights.

Implications

The current study, along with previous work (Jarvis & Harris 2008), shows that the high accident rate of inexperienced UK glider pilots is not caused simply by more accidents in general occurring to these pilots, or even a higher frequency of the sorts of accidents occurring to other glider pilots. The vast majority of crashes for low-hours pilots are accounted for by specific kinds of accidents in specific phases. Indeed without these accidents, the accident rate for early solo pilots would be relatively low! This prompts important questions regarding the release of student pilots for solo flight in the first place.

The following chapter addresses whether instructors underestimate the risk faced by inexperienced pilots, particularly in these flight phases, and begins to address questions relating to instructor judgement in allowing solo flight.

Chapter 7 - Further Technical Background

Instruction and supervision of trainee glider pilots in UK gliding clubs

As stated in the previous chapter, there is a need to focus on the nature of instructor judgements with regard to solo flight. Because subsequent studies will begin to explore these issues there is a need to present a small amount of essential background concerning post-solo supervision and the nature of UK gliding instruction.

Post solo restrictions

All gliding clubs operate some form of restrictions by which minimum pilot qualifications or experience are required in order to fly when certain conditions prevail as judged by the instructor in charge (notably weather conditions, but may include ground conditions, anticipated conditions, low sun, etc). Many of the bigger clubs operate a colour system with variously coloured windsocks hoisted dependent upon the duty instructor's assessment of flying conditions. Some clubs issue flying cards to members based on experience and qualifications. For example at Lasham airfield (the largest gliding centre in the UK) white, red, yellow and blue cards are issued as pilots pass a laid down criteria. The colour of the windsock directly relates to this colour coding, meaning that only pilots holding that colour card or above may fly solo.

As well as restricting early solo flying to safe flying conditions, clubs restrict solo flying based on currency (the amount of recent flying). The BGA states "Each club should develop, publish and implement a set of currency requirements for all pilots" (British Gliding Association 2006c). No specific rules are laid down by the BGA, but advice is given on how to decide. The first piece of advice is that pilot experience, in terms of "hours, launches, number of types flown, number of sites etc", should be taken into account (British Gliding Association 2006c).

In addition to restrictions based on conditions and currency, inexperienced pilots are only allowed to fly certain aircraft types such as training gliders, and are only allowed to fly within range of the gliding club.

Such restrictions allow pilots to build up experience before facing more difficult situations caused by poor weather, flights after a small lay-off, or field landings.

Solo Supervision – Differences between gliding and powered flying

A big difference between glider training and powered aircraft training is the approach taken after the initial solo flights.

The level of supervision for powered pilots after their first solo flights remain high until they have their PPL. Indeed the pilots are still seen as trainees who still do the majority of their flying with their instructor and are only allowed to fly solo under strict supervision.

The approach taken in gliding has notable differences. ‘Solo supervision’ is practiced meaning that before flying on a given day or a given runway direction, early solo pilots may receive one or more ‘check flights’ with an instructor who decides whether they are competent to fly solo in those conditions. Such flights often include some sort of simulated emergency situation. However these flights are not usually part of a training syllabus, but assessment flights by the instructor to make sure the pilot is competent to fly on that day.

Glider pilots soon begin a process of being signed as ‘off-checks’. Many clubs have a formal system. For example The London Gliding Club has a set of formally tested goals prior to the bronze badge called ‘Performance Assessment 1’ (PA1) and ‘Performance Assessment 2’ (PA2). Until a pilot has passed their PA1, they must receive a check flight with an instructor prior to flying on any given day. However this also means that having passed the PA1, a pre-bronze pilot is officially allowed to fly solo without briefings or check fights.

It is normal for this 'off-checks' stage to occur well before the Bronze badge is achieved because systems such as PA1, PA2 or coloured cards come into play prior to the Bronze badge. Hence, unlike pre-PPL pilots pre-Bronze badge glider pilots can arrive at their gliding site and fly without the need for briefings and check flights, as long as they are within their own set of restrictions for the day. Hence the new solo glider pilot's focus is solo flying, unlike the new solo power pilot.

The BGA instructors' manual recognises the problem: "When trainees reach the 'off-checks' stage it doesn't mean that your interest in their progress should cease abruptly. Their conduct will still need monitoring" (British Gliding Association 2003). However there is no formal system for this monitoring.

The big difference from powered aircraft pilots is that prior to the issue of the PPL, powered aircraft pilots remain inside a very structured training program, whereas structured training all but stops for UK glider pilots once they fly solo, and particularly once they are 'off-checks' (which could follow quite soon afterwards).

Instructor categories

Gliding instructors in the UK are appointed by the British Gliding Association (BGA), and operate under the Chief Flying Instructor (CFI) at a given club. In order to become an instructor, a pilot must have 50 hours in command and a FAI Silver Badge. He or she must then train at their home club, usually with senior instructors appointed by the CFI. Once ready the pilot attends an instructor course run by the BGA. The first instructor rating is known as Basic-Instructor rating and requires a weekend course. This rating allows the new instructor to instruct gliding only above 500ft above ground level (AGL), which in practice means that 'basic instructors' only teach introductory gliding lessons. The second rating is the Assistant Instructor rating. This involves a nine-day course and further two-day completion course. This rating allows the instructor to instruct student pilots all the way to solo and beyond. Lastly is the Full-rating. This involves logging experience as an assistant instructor and gaining approval after a course with a senior BGA coach. In practice, and from the point of view of a student, there is almost no difference between assistant and full rated instructors. The only official difference is that CFIs must have a full rating.

Instruction of UK glider pilots

Instructional practice is outlined in the BGA instructors' manual, which is updated and revised by the BGA instructor committee. Detail is given on how all flight phases should be taught, and some specific exercises are included such as particular stall and spin demonstrations, lookout exercises, etc. On the other hand little advice is given on the order of exercises or student progress, or how to decide if a student is ready to fly solo. Some of this will be looked at in the next study.

Gliding is a recreational activity, with very few paid staff in the UK. Several of the larger clubs employ a professional chief flying instructor, and sometimes a staff instructor during the season to run courses, but volunteer instructors do the majority of club member instruction. The vast majority of clubs in the UK run on an entirely voluntary basis, with instructors given a roster of their duty days, typically once a month depending on the number of volunteers. For this reason most pilots learn to glide over a period of months, flying with numerous different instructors, some of whom they may meet for the first time prior to the flights. This is different to most forms of aviation and has possible unrecognised implications. General aviation instructors are able to build up a picture of how the student is progressing through their training, since student pilots do most of their flying with just one or two instructors. In UK gliding this rarely happens. Instructors are kept informed of the student progress by conversation with the student, looking at the logbook, checking progress cards (if the club uses them), and any casual conversation with other instructors who have flown with the student previously. In gliding therefore student progress is rarely formally monitored. This means that the critical decision to send a student pilot solo is often taken by an instructor who has not had the chance to build up a picture of the student's progress through his or her own experience. They may not even have flown previously with the student.

Chapter 8 - Investigation into instructors' perceptions of the relative likelihood of accidents among early-solo pilots and the relative danger presented by each flight phase

Accident rates have been shown to be disproportionately high amongst early-solo glider pilots (Jarvis & Harris 2007b), and the vast majority of these accidents are pilot-related (Jarvis & Harris 2008). As well as the accident rates being high for this group, it is now known that this is not caused simply by these pilots having more accidents in general, or even a higher frequency of the sorts of accidents occurring to other glider pilots. Low-hours pilots have specific kinds of accidents in specific phases (namely, accidents caused by misjudgements and mishandling during the final approach or landing flare). These accident causes are relatively unparalleled in the rest of the UK glider pilot population (see chapters five and six).

The high accident rate to low hours pilots occurs against a background of post-solo supervision which includes check flights, restrictions (only flying benign glider types in benign conditions) and limitations on flying range. Hence these pilots have the highest accident rate despite being most likely to fly only in the most favourable circumstances. It was previously noted that the picture that emerged from chapters one and two suggested that pilots learn rapidly once solo and soon enter a safer experience bracket. If they can achieve over ten hours solo then their chance of having an accident diminishes rapidly. All of this must focus attention on the instructors, since all very inexperienced pilots must have recently been judged safe to fly alone by a qualified instructor. Instructors may be allowing pilots to fly on their own before some of their most important learning is completed, particularly in specific areas (approach and landing), and unknowingly exposing them to a high level of danger for at least their first few hours flying.

Given claims made in popular aviation literature regarding experience and accident rates (see Chapter 2), it is possible that there is a general underestimation of the dangers faced by low hours-pilots within the UK gliding movement. Many authors (of research literature and popular textbooks across aviation) claim that very inexperienced pilots are a relatively safe group (Booze 1977, Piggott 1997, Pratt 2000, Telfer 1993). There are also many claims that particular experience levels are associated with high accident involvement. These are generally stated to be between 100 and 500 hours (Olsen & Rasmussen 1989, Jenson 1995, Wells 1992). The reasons often cited for these phenomena are centred on terms such as 'risk taking behaviour', 'overconfidence' and an increased perception of 'invulnerability' (Booze 1977, Telfer 1989, O'Hare 1990, See Chapter 2). Such terms as applied to aviation accidents have been criticised as "folk models" because they are easily applied labels but do not explain accident causes (Dekker 2006). Looked at carefully these explanations can be seen to be related to the pilot's attitude (i.e. a predisposition of the pilot) rather than the situational circumstances. Theories of psychological attribution describe the tendency of individuals to over-estimate the role of dispositional factors (in this case pre-existing pilot attitudes), a bias referred to as *fundamental attribution bias* (Ross & Anderson 1982). It has since been shown that people judge dispositional characteristics before correcting their judgements on the basis of situational characteristics, the latter process requiring greater mental effort (Gilbert 2002). With limited information and anecdotal evidence on which to make such 'corrections' to initial judgments it is not difficult to see why people might conclude that accidents were caused by the predisposition of the pilot (attitudes, abilities, etc).

Given the claims made over the last 30 years proclaiming the relative safety of inexperienced pilots and high accident rate of more experienced pilots, it is possible that the issue of instructor judgement (in allowing solo flight) has been overlooked because it was not regarded as a priority. If a common belief exists that accident vulnerability increases with pilot experience (albeit to a specific level prior to declining again) then one can be forgiven for thinking that instructor judgement of students is already effective, or is not important, and the focus should be on more experienced pilots. Indeed, many clubs run involved annual check flight sessions in which pilots of higher experience levels are checked for competence by instructors. It is possible that instructors, and the gliding movement as a whole, pay too little attention to very inexperienced pilots as a result of popular misunderstandings about their relative level of vulnerability.

This study asks whether instructors underestimate the relative accident probability of pilots that they send solo (for their first ten hours), particularly in the approach and landing phases, in order to start answering questions relating to instructor judgement in allowing solo flight to take place.

The British Gliding Association is a very small body with few professional staff. Gliding instruction is provided at club level mainly by volunteer instructors. This is a very varied group, but makes up the majority of the influential membership in the British Gliding Association, in terms of training and instruction. Many instructors are also examiners, chief flying instructors, instructor trainers, or just highly experienced instructors operating across several clubs. Hence as well as being responsible for allowing pilots to fly solo, club instructors make up the major body of expert knowledge in the UK gliding movement, and are therefore appropriate as the target population for this study.

Since it is known that expert evaluations of the probability of uncertain events can play important roles in decisions (Tversky & Koehler 1994), it is probable that instructors' perception of the accident likelihood faced by newly-soloed pilots affects decision making at various levels and even policy regarding new solo pilots. It is commonly recognised that "risks are often produced when dangers are overlooked or underestimated" (Sjöberg 1987). Therefore eliciting information regarding instructors' general perception of the accident likelihood faced by glider pilots as well as the parts of the flight that cause most problems will start to build a picture of the situation from the instructors' perspective.

The decision to send a pilot solo is done without advanced knowledge of the consequences, and any such decision is a 'risky' choice (Kahneman & Tversky 2000). The concept of 'risk' is not easily defined with reference to literature, since so many different definitions exist and there is no common understanding of the term (Renn

1998). Many definitions rely on mathematical calculations of numerical data and hence appear objective in their interpretation. A classic 'engineering' definition is "a combination of the probability, or frequency, of an occurrence of a defined hazard and the magnitude of the consequences of the occurrence" (The Royal Society Study Group 1992). However such apparent objectivity is seriously questioned, particularly within the human science domain; some see all risk as subjective; an invention by human beings to help understand and cope with dangers and uncertainty, meaning there is no such thing as real-risk (Slovic 2000). Clearly, for the purposes of this study, a workable concept, or definition was required.

Classic definitions, such as that of The Royal Society (1992) see 'risk' as a dependent variable, defined by a combination of two independent variables, 'consequence' and 'probability'. In terms consequence it has been shown from accident data that there is no statistical difference between low-hours glider pilots (10 hours and under) and other glider pilots for the five years 2002 - 2006 inclusive (see Chapter 4). Fisher's exact tests showed no significant association between pilot experience groups and the degree of injury ($p=0.701$) or glider damage ($p=0.272$). However, while accident consequences show no statistical difference, the probability of low-hours pilots having an accident was very much higher than for other glider pilots (Jarvis & Harris 2007b). Hence this group are at just as much 'risk' in terms of consequences when an accident occurs, but far more likely to have an accident in the first place. This means that in terms of 'risk', the independent variable of most relative impact for this group (relative to other glider pilots) is accident likelihood, not consequence severity.

Because of this, this study does not concentrate on 'risk', but on accident likelihood only. Although some authors have simplified the use of the word 'risk' to this degree, e.g. "the probability of an unfavourable outcome" (Patankar & Taylor 2004), the term 'risk' was not used as the operational concept in the present study. This was because the intention was not to measure risk itself or the perception risk, but simply to discover whether instructors' views regarding the relative likelihood of accidents aligned with the research from accident data. However, since this involved elements of 'risk', some literature pertaining to risk was reviewed.

Since the study aimed to establish instructor perceptions of accident rates, two components were required; a 'real' measure, and a perceived measure. Studies of risk-perception often compare objective measures of probability to the perception of risks (Bohm & Harris 2009a). Conventionally in risk studies, objective measures are deduced from rates based on data of past accidents (Bohm & Harris 2009a). Previous studies (Jarvis & Harris 2007b, 2008) provide such data as well as rate calculations for relative accident rates of inexperienced pilots and flight phase causality. This will make it possible to determine whether instructors' perceptions of relative accident probability of newly solo pilots are accurate. Many studies have observed that workers' risk perception was significantly different to measures of 'objective' risk (Bohm & Harris 2009a).

The accident data show that the probability of an early solo glider pilot having an accident during their first hour of solo flight is approximately one in 713 (see Table 2.3, chapter 2). It is unlikely that such a figure would be meaningful to a gliding instructor when quantifying likelihood (or making solo decisions) since no equivalent figures have ever been produced.

For reasons of this kind, most studies of subjective risk assessment require individuals to assess the risk of one thing against the risk of another, rather than to give absolute values (e.g. Bohm & Harris 2009a, Weyman & Clarke 2003, Ostberg 1980). This is often in the form of a ranking task, and makes the process more meaningful to the participant. Hence, to assess general accident probability for inexperienced pilots, a comparison will need to be made to more experienced pilots. To assess the accident probability associated with a particular flight phase is straightforward because each flight phase can be ranked against each other flight phase. The accuracy of the instructors' perception can then be found by comparing their evaluations with comparable information from the previous studies on accident data. Hence this study aimed to elicit instructors' level of appreciation of the situation (regarding pilots with ten hours or fewer) as defined by previous work on accident rates and flight phase causal factors (Jarvis & Harris 2007b, 2008).

When evaluating probabilities in relation to groups or populations with unequal frequencies (such as inexperienced pilots within the larger population of all pilots) it has been shown that there is a tendency to neglect the underlying difference in the sizes of the groups. This has been termed 'the neglect of base rates' (Kahneman et al 1982). Base rate neglect is a common human bias, as described by Hogarth (1987);

> "Probability theory argues that one should modify base rates by case data such that the ensuing judgement reflects both [specific and base-rate data]. People's intuitions, however, do not always correspond to the laws of probability in this instance. Indeed considerable evidence has been documented of the failure to consider base rate data."

Neglecting to take into account the relatively small size of the population of inexperienced pilots when reflecting on accident figures or anecdotes would have the effect of obscuring the high accident rates sustained by the smaller group, and hence could lead instructors to underestimate the relative danger (in terms of accident likelihood) faced by newly soloed pilots.

It is quite possible that instructors' perception of accident probability impacts on the decision making process when allowing such flights to take place, since decision making in situations of risk can be viewed as a choice between prospects and gambles (Kahneman & Tversky 1979). One of the best known models of decision making under uncertainty is 'expected utility theory' also known as 'decision theory', which postulates that people select the alternative with the greatest expected utility (simplistically the most favourable outcome) using rational methods (Hogarth 1987). One of the best known generalisations of the theory is that people are risk averse (Kahneman & Tversky 1979). On the other hand, 'prospect theory' (Kahneman & Tversky 1979) shows that this is only generally true when people face gains, but in the face of losses many people become more likely to take risks. The instructor decision to send a pilot solo has only one option that involves any level of risk; allowing solo flight, and so it is not straightforward to determine where the losses and gains would factor into the decision for the instructor. Both of the stated theories would appear to predict that instructors avoid the risk of allowing solo flight (risk aversion), since there is no apparent need to allow the flight and there are no obvious losses to be faced. Clearly however, at some point, instructors do take the risk and send the pilot to fly alone. There may be other factors involved which are viewed as losses or gains such as possible strain on relationships (in not sending someone solo), or gain of credibility and praise (for doing so).

As well as attempting to assess potential losses and gains, in the face of uncertainty individuals tend to rely on heuristics, which are speculative frameworks used to guide solutions or understand the world (Botterill & Mazur 2004). Estimation of probability of undesirable outcomes has been shown to be biased by several heuristics, of which the 'availability heuristic' is generally agreed to be of most importance in understanding risk perception (Botterill & Mazur 2004, Slovic et al 1982). This is a process by which people judge an event as likely or frequent if instances of it are easy to imagine or recall (Slovic et al 1982). Amos Tversky and Daniel Kahneman first recognised the existence of this heuristic, and define it as "the ease with which instances or occurrences can be brought to mind... a useful clue for assessing frequency or probability" (Kahneman et al 1982). The availability heuristic manifests itself is as follows: Members of one class (e.g. population of people, set of objects, etc) will be judged as more numerous than another equally sized class if its members are more 'retrievable' in terms of memory. An elementary example would be a list of names containing equal numbers of men and women, but containing more famous men than famous women. The availability (in terms of memory retrieval) of the famous men will make men appear more numerous than women (Kahneman et al 1982). In this case the availability would be caused by familiarly, but it can also be caused by the salience, for example witnessing real events rather than seeing photographs (Kahneman et al 1982). In terms of decisions to allow solo flight, availability bias might lead to instructors failing to recognise the high accident rate. As previously stated, the vast majority of accidents happen to pilots with more than ten hours (simply because there are many more such pilots in the population). Availability bias means that people tend to think that events are more probable if they can recall an incident of their occurrence (Botterill & Mazur 2004), and with the very few total accidents occurring to newly soloed pilots, most instructors would be unlikely to recall as many cases, if any. They would however be likely to recall cases of accidents occurring to more experienced pilots. As well as being more numerous, these accidents are more likely to be the sort of accidents instructors might have while flying on their own (e.g. field landing accidents) and so make more of an impression on instructors and hence be more salient in memory.

A similar phenomenon that has been shown to elevate people's estimation of the probabilities of certain events is the feeling of 'perceived dread' associated with outcomes (Slovic 2000, Bohm & Harris 2009a). This would appear to be related to salience as previously mentioned, and has been shown to affect evaluations of environmental hazards such as nuclear accidents (Slovic 2000) as well as the perception of accidents in the workplace (Bohm & Harris 2009a). It is possible that such 'perceived dread' does not exist for instructors when deciding to send a pilot solo because there is no possibility of 'dreaded' consequences for the instructor, and hence this may result in an underestimation of accident probability. Indeed this is supported by a finding that SMEs in the construction industry were not affected by feelings of perceived dread in the way that the equipment operators themselves were (Bohm & Harris 2009a). Perceived dread would make it more likely that instructors base their assessment of the dangers of early solo flights on the aspects that they themselves fear most, or that they fear their students doing while they are with them in the aircraft.

The decision to allow a new pilot to fly solo has huge safety implications. From previous studies (Jarvis & Harris 2007b, 2008) there is sufficient evidence to suspect that glider pilots are often sent solo before they are ready. Since a large amount of aviation literature underestimates the danger faced by early solo pilots (Jarvis & Harris 2007b) it is very possible that the UK gliding movement (in the form of its instructor core) underestimate the likelihood that the people they send solo will have an accident. This could bring the question of judgement in decision making for early solo flying into question.

No research has attempted to explore risk perception, or the relative accident likelihood perceived by flying instructors in relation to their students, nor the effect of this on the judgement itself. The first aim of this study was to discover the extent to which gliding instructors are aware of the level of vulnerability of pilots recently sent solo (relative to other pilots), and the reasons behind their estimates. The weight of popular assertions in aviation literature and the influence of cognitive biases such as the availability heuristic and base rate neglect, lead to an expectation that instructors will underestimate the relative likelihood of early solo pilots being involved in accidents, compared to more experienced pilots. It is more likely that instructors will estimate the pilot experience level associated with the highest accident involvement as being between 100 and 500 hours, and base these estimations on pilot predispositions more than the situational characteristics in line with claims in the literature.

The second aim was to determine whether instructors were aware of the parts of the flight that caused most accidents to this vulnerable group. This may be affected by factors such as the 'dread factor' and the availability heuristic. This would involve probabilities being biased by aspects of the flight that the instructors are most afraid of, the most salient situations they recall, or those parts of the flight in which they most recall students making errors during instruction. It has been shown that in terms of instructional flights, the launch and landing stand out as being the phases that cause a disproportionately high number of accidents (see Chapter 5, Table 5.10).

Method

Procedure

A questionnaire was distributed to gliding instructors that requested them to indicate what pilot experience levels they believed were most associated with accidents, and give a short explanation for their figure. Additionally participants were asked to rank six flight phases in the order that they thought led to most accidents.

Sample

For sound practical reasons, all data collection was completed promptly. Notwithstanding this it was important to sample instructors from a number of different clubs and of different experience levels and ages. Two gliding clubs were chosen as the main centres for data collection. One was a large club focussing mainly on training and cross country soaring, the other a medium sized club focussing mainly on general soaring and glider aerobatics. Just over a quarter of the participants were approached individually at other UK clubs.

Instrument

The questionnaire (Appendix E), contained three sections; participant demographic information, estimation and explanation of pilot experience most associated with accidents, and flight phase ranking. After the demographic section, instructors were asked to mark a horizontal scale from 0 to 5,000 hours to indicate their estimate, and then write the exact figure in the box provided. It was explained that they could make several estimates if they felt that there were several different pilot experienced levels associated with high accident rates. If so, they were asked to circle the one that they felt was associated with the highest accident rate. Participants were instructed that if they perceived there to be a bracket of pilot experience associated with high accident rates, then they should indicate only the lowest point of that bracket (e.g. a bracket between 200 and 500 hours would simple require them to write '200' hours). After this participants were asked to explain as briefly as possible (in one sentence) their perceived reasons for the high accident rates at the experience level(s) that they had identified. Participants were fully briefed as well as being given written instructions (Appendix E). The instructions were written and agreed upon by the investigator and two other gliding-related SMEs. Further validation involved reading the instructions to a small sample of gliding instructors. The feedback showed that the meaning was clear and as intended. In the final part of the questionnaire, participants were presented with six flight phases (pre-flight, launch, top of launch to circuit, circuit, approach, landing) in line with the task break down of Jarvis & Harris (2008). They were instructed to rank these from six to one (in the boxes provided) to indicate their perception of the relative likelihood of each phase to cause accidents to pilots with ten hours or fewer. They were instructed that six indicated the flight phase that they thought was associated with the most causal accident events and one for the least. In order to prevent bias caused by the ordering on the questionnaire, the flight phases were presented in a circle rather than a list, and six batches of the questionnaires were produced with them ordered differently in each. These were shuffled together prior to distribution.

Data collection

Two specially arranged instructor evening presentations were arranged (one at each participating club, run by the investigator and organised by the clubs), in which questionnaires could be completed and returned in large numbers. Over two thirds of the questionnaires were filled in at the two special sessions. This prevented the chance of communication occurring within each club and distorting the results. At the meetings, participants were briefed collectively before the questionnaires were distributed. Silence and discretion were requested. After collection of the questionnaires participants were de-briefed as part of the presentation, and results of Jarvis and Harris (2007a) as well as Jarvis and Harris (2007b) were presented as part of the evening's presentation. The remaining data collection (run on a convenience basis at a variety of clubs due to the volunteer nature of UK gliding) was conducted by the investigator in person and the same briefing was given.

Demographic data were collected from all participants in order to check the fairness of the sample.

Treatment of data

Quantitative data.

For the analysis of the flight phase data, the method and data analysis drew upon the work of Bohm and Harris (2009) who determined risk perception of seven hazardous scenarios by dumper drivers on construction sites. An objective ranking of the six flight phases was prepared from accident data using calculated accident frequencies (from rates) for pilots with 10 hours or fewer (from Jarvis & Harris 2008). This objective ranking was then compared to the subjective ranking of the participants, which had been calculated by obtaining a mean rank for each flight phase and then ranking the resulting means.

Qualitative data.

Each narrative was treated as an individual unit of meaning in its entirety, since none were very long. The paired-comparison method as described by Partington (2002) and previously employed in Chapter six was used to compare each statement with its preceding statement to see if the same phenomena were described. Where this was the case they were coded as belonging to the same category. After the initial process was complete, the categories were named and all statements re-coded into those categories. This process required further refinements to the categorisation and labelling. This process was repeated a number of times until a satisfactory set of representative categories were established that adequately described the themes within the data.

Results

Questionnaire - part 1: Demographics

There were 74 participants, consisting of 34 full rated and 40 assistant rated BGA qualified instructors, 72 were male and two female. The Mean gliding experience was 1251 hours (sd = 1103) and 20.6 years (sd = 11.1). Age was requested in ten year brackets, with the distribution as follows (n represents the number of participants in that bracket);

[20 - 29]	n = 2
[30 - 39]	n = 7
[40 - 49]	n = 18
[50 - 59]	n = 23
[60 - 69]	n = 17
[70 - 79]	n = 7

Questionnaire - part 2:

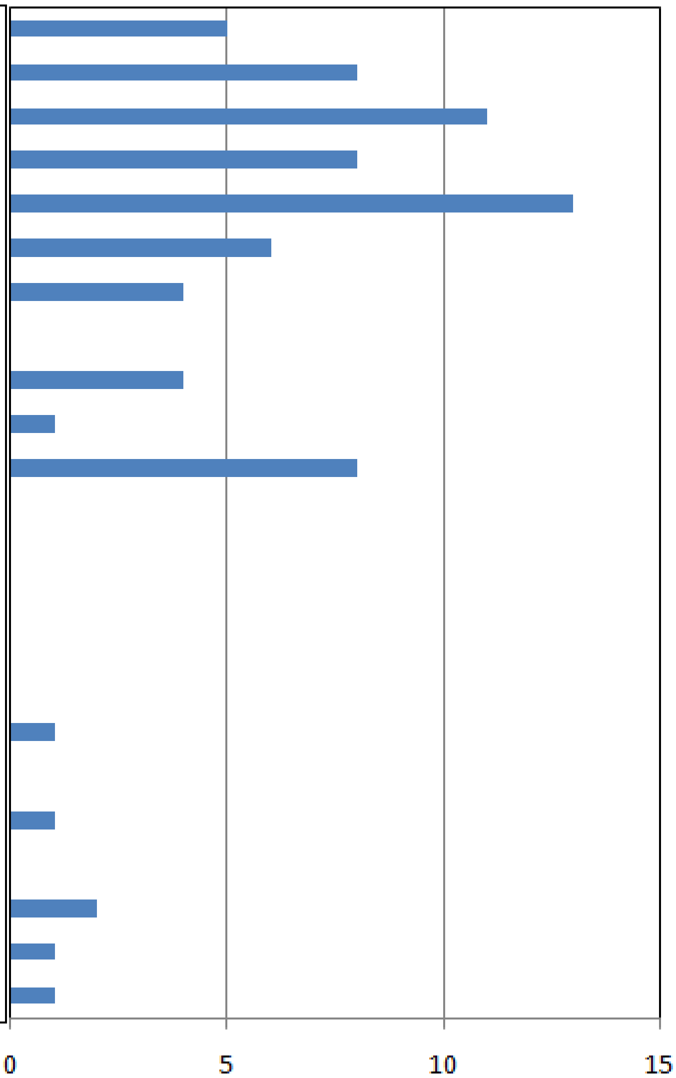
A. Estimation of the pilot experience levels with high accident rates

Only seven participants indicated a second or third pilot experience level (or bracket) as bring was associated with relatively high accidents rate for glider pilots. For each of these only the level associated with highest accident rate (as they indicated) was used in the analysis.

The overall estimation of peak accident experience was 288.7 flying hours (sd = 314.2). Table 8.1 shows that a large number of participants considered that there was an increase in accidents between 100 and 250 hours (n = 37, exactly half the participants). Fifty-five participants (74%) estimated a figure between 100 and 500 hours inclusive. Only 13 participants (18%) estimated a figure below 100 hours, and only one put the figure within the vulnerable 10 hour bracket (at 10 hours) which included all the data from those giving multiple estimates.

Table 8.1. The percentage (%) and number (n) of estimates in each 50 hours experience bracket. The accompanying bar chart illustrates the frequency (n) as represented by the horizontal axis. For example, the top row shows that five participants (6.76%) gave an estimate between 0 and 49 hours as the level of pilot experience they thought was associated with the highest accident rate.

Estimate(h)	%	n
0 - 49	6.76	5
50 - 99	10.81	8
100 - 149	14.86	11
150 - 199	10.81	8
200 - 249	17.57	13
250 - 299	8.11	6
300 - 349	5.41	4
350 - 399	0.00	0
400 - 449	5.41	4
450 - 499	1.35	1
500 - 549	10.81	8
550 - 599	0.00	0
600 - 649	0.00	0
650 - 699	0.00	0
700 - 749	0.00	0
750 - 799	0.00	0
800 - 849	1.35	1
850 - 899	0.00	0
900 -949	1.35	1
950 - 999	0.00	0
1000	2.70	2
1200	1.35	1
2000	1.35	1



B. Qualitative explanations given by participants for their estimated figures

Seven categories were established, three of which were deemed to describe a single overlying concept (pilot attitude), and were hence merged into one large category (category 1). The category titles and the numbers of explanations coded into each were:

1. Attitude (n = 38)

- **1.1.** Over confidence / complacency (n = 23)
- **1.2.** Taking risks / Pushing boundaries, under-estimating dangers (n = 12)
- **1.3.** Enthusiasm / Ambition (n = 3)

2. More challenging types of flying, conditions, gliders, situations (n = 15)
3. Post solo supervision related (n = 11)
4. Known Figure (n = 6)
5. Reason given for improvement (n = 4)

Reliability check

Reliability was checked independently by a second rater (a human factors professional with research experience) who re-coded all 74 explanations into these seven established categories (including all three 'attitude categories'). The reliability matrix is shown in Appendix F. Sixty-five out of 74 statements were coded the same as the original rater, giving a percentage agreement of 88% (Cohen's Kappa .848). The remaining nine statements were discussed and agreement reached in all cases about which category to code them by and why.

Descriptions of each category and example statements

1. Attitude: This was a large category formed by the merger of three others, since they were all formed of comments relating to pilot attitude. This category contained just over half of all comments made.

1.1. Over confidence / complacency: This category was formed as a response to the number of comments using the terms 'confidence' or 'complacency' or both. Only responses using these terms or obviously referring to them were placed in this category. It reflects the expression by instructors of an undesirable attitude demonstrated by pilots. Nearly a third of all comments were to do with confidence or complacency.

Examples:

- "Confidence outstripping capability and getting over-complacent"
- "Young and over-confident, 'Nothing will get me' attitude"
- "Confidence levels are at a max until about 200 or 300 hours"

1.2. Taking risks / Pushing boundaries, under-estimating dangers: This reflected statements concerned with the notion that pilots increasingly start taking more risks or 'pushing the boundaries'. Occasionally this was linked to competition flying.

Examples:

- "Risky behaviour reaching its peak"
- "They start to feel immortal and start taking risks, after a few scares settle down more, if their (sic!) still alive"
- "a peak risk seeking time, pilots can't assess the risks accurately because inexperienced"

1.3. Enthusiasm / Ambition: This category represents just a few comments made suggesting that pilots had accidents precipitated by their level of enthusiasm or ambition.

Examples:

- "First few solos are safest then they get over enthusiastic, bust season, trying everything"
- "over-keen"

2. Doing (or encountering) more challenging types of flying, conditions, gliders, situations: This category reflected many comments about high accident rates being related to the increased challenges encountered or engineered with experience. These included flying further from the base airfield, cross-country flying and competitions. Other increased challenges were mentioned to do with conditions, more challenging kinds of flying, and more advanced gliders.

Examples:

- "Flying further out of the local area and first few field landings are a big danger, then competitions leading to lots of broken fibreglass."
- "Season three or 4, first serious cross country flights and badge attempts"
- "people buying into advanced and slippery machines and doing longer flights"

3. **Post solo supervision related:** This was the idea that accidents were precipitated due to the diminishing supervision that accompanies increasing experience. Since a brief statement was asked for, most do not offer more detail as to why this occurs.

Examples:

- "Pilots are protected in their first few hours (we keep an eye out etc) hence they are safest then, but they become freer to do the things that eventually leads to the inevitable"
- "They are stepping out of the post-solo environment, post solo supervision, no one looking out for them and keeping them in line"
- "Post solo is so poorly looked after and trained in the UK, lots of accidents once pilots are left to fend for themselves, off-checks."
- "By 50 hrs they are getting away from the heavily regulated part of their gliding career which kept them safe"

4. **Known Figure:** This category included all comments where participants stated that their indicated figure was previously known.

Examples:

- "well known figure, between 200 - 400"
- "Statistics ?? 400 to 600 hrs. From the RAF originally"
- "Common knowledge of the statistics (sorry!) 250 to 500 is the best known (from the BGA?)"

5. **Reason given for improvement:** Some participants described why pilots became safer after the peak / zone, as opposed to more vulnerable as they entered it. In these cases the comments were entered into this category.

Examples:

- "The feeling of reality hits and you don't take as many risks, So things get better after this (for most people!)"
- "Enough scares to realise this thing's going to kill me sooner or later"
- "Reduction in risks as they know what can happen"

The analysis suggested that three broad causes were given by instructors as to why accident rates should increase sometime after going solo. These were

- Caused by an increasingly unsafe attitude (category 1)
- Caused by an increasing level of challenge (category 2)
- Caused by a reduction of post solo supervision (category 3)

The other two categories did not contain suggestions relating to the causes of the accident rates.

Effect of instructor qualification, age and experience.

Pearson's chi-square was used to show that there were no significant differences between full and assistant rated instructors for frequencies of categories. For example the number of 'attitude related' comments (category 1) was 20 out of 40 for assistant rated instructors and 18 out of 34 for full category instructors. This results in a chi-square of 0.064 (df = 1) which is not significant (p = 0.1).

There were no significant differences between instructors on the basis of their own total flying experience or age. Ten out of 19 instructors with 500 hours or fewer gave reasons relating to attitude, 18 out of 31 with between 501 and 1000 hours did so, and 10 out of 24 with over 1000 hours. This relates to a chi-square of 1.473 (df = 2) which is not significant (p = 0.1). In terms of age, 13 out of 27 of those under 50 gave reasons to do with pilot attitude, as opposed to 25 out of 47 of older instructors (50 plus). This results in a chi-square of 0.175 (df = 1) which is not significant ($p > 0.05$).

Participants' estimated figure (for the pilot experience level associated with the accident rates in UK gliding) against the reasons given for it.

Table 8.2. Mean participant estimates of experience level for highest accident rate, associated with statements given in each of the qualitative categories. I.e. Each mean figure represents only the estimates of instructors who gave explanations coded as belonging to that specific category.

	Category	**Participant mean estimate (hrs) of experience level for highest accident rate**	**sd**
1	**Developing an unsafe attitude**	306	238.2
2	**Exposure to greater challenge**	402	546.0
3	**Post solo supervision related**	66	41.2
4	**Known Figure**	330	141.8
5	**Reason given for improvement**	250	129.1
1.1	Over confidence / complacency	331	283.9
1.2	Taking risks / Pushing boundaries	282	143.1
1.3	Enthusiasm / Ambition	213	162.9

Table 8.2 shows that participant estimates of peak accident rates that were accompanied by explanations relating to post-solo supervision (category 3) were, on average, a lot lower than the other groups (only 66 hours experience). Explanations relating to pilots being exposed to greater challenges were accompanied by the highest estimates (mean of 402 hours).

3. Flight phase data.

Fifty-eight of the participants filled in the flight phase questionnaire. Table 8.3 shows the accident likelihood for pilots with ten hours or fewer experience (the objective measure) produced from findings of Jarvis and Harris (2008), against rankings produced from the subjective instructor assessments ('6' means the highest accident likelihood and '1' the lowest). Comparing the subjective ranking against the objective ranking shows that two out of the six flight phases match ('Top of launch to circuit' and 'Landing') and two are within one ranking position (pre-flight and circuit). However 'launch' and 'approach' are ranked too high and too low respectively. From the accident rates the launch has a low frequency of causal occurrence (second lowest flight phase occurrence) whereas the subjective judgements suggest that instructors believe it to be the highest causal phase (ranked '6'). The approach phase, despite being the highest accident causation flight phase for low-hours pilots was ranked as second lowest by instructors.

Table 8.3. Objective ranking based on the number of flights per accident associated with each flight phase (from Jarvis & Harris 2008) against subjective probability (instructor perception) of these rankings (from the mean ranking for each flight phase from the 58 questionnaires).

	Objective Accident Probability (from accident data) (10 hrs and under)		Subjective Accident Probability (instructor perception) (10 hrs and under)	
Flight Phase	Launches per accident	Rank	Mean	Rank
Pre-flight	9975	3	3.40	4
Launch	14962	2	4.71	6
Top of launch to circuit	29924	1	2.52	1
Circuit	9194	4	3.31	3
Approach	2720	6	2.71	2
Landing	3741	5	4.38	5

Discussion

Estimation of experience level and accident involvement

As expected the perceptions of UK gliding instructors closely reflects the literature. In common with claims of low hours pilots being safe (Booze 1977, Pratt 2000, Telfer 1993), only 13 instructors (18%) estimated peak accident rates to be associated with pilots with fewer than 100 hours, and only five (7%) put a figure less than 50 hours (Table 8.1). No instructors indicated that they believed pilots with ten hours or fewer had a relatively high accident rate and this held true even when the additional estimates of those participants who had multiple estimates were taken into account. Some explanations offered went as far as to state that early-solo pilots were safest (e.g. "First solo is pretty safe then peaks and troughs", "First few solos are safest"). Given the trend of accident rates discovered by Jarvis and Harris (2007b) this means that in common with Piggott (1997), instructors seriously underestimate the relative danger faced by UK glider pilots embarking on their first few solo flights.

The implications of this could extend into instructor decision making, particularly given that instructors must decide when it is safe to allow a student pilot to fly solo. It was noted that decision making under uncertainty involves some assessment of probabilities and choices between prospects and gambles (Kahneman & Tversky 1979). Major theories (expected utility theory, prospect theory) involve some estimation of losses, and if an instructor underestimates the chance of a loss (undesirable outcome) this could lead to students being sent solo before they are safe to do so. However, very little is known about instructor decision processes when sending a pilot solo.

The finding that most participants (74%) estimated a figure of between 100 and 500 hours generally reflects the claims of similar figures in the literature (Olsen & Rasmussen 1989, Jenson 1995, Wells 1992 & Telfer 1989). Participants may have been influenced by reading the literature, the literature may have been influenced by perceptions held across the aviation community, or both the community and the literature may have misread the situation in the same way, possibly for the same reasons.

There is some evidence suggesting that instructors may have been influenced by the literature. The qualitative category labelled ‘known figure’ (see ‘Results’ section B), contained some comments clearly stating that the participant had read of the figure previously, and additionally all of these stated (as text) that it was a bracket (variously between 200 and 600 hours) which is very similar to claims in the literature. This does suggest that the literature has had an effect on perception, at least for some.

However, an explanation given by Pratt (2000) for a claim regarding the relative safety of student pilots indicates a possible common reason why the literature and the community alike may be underestimating the relative probability of accidents occurring to early solo pilots. Pratt states that “less than 5% of the fatalities were student pilots; far more instructors and professional pilots were involved in fatal accidents”. Just as it has been pointed out that many studies fail to adequately control for levels of exposure (Jarvis & Harris 2007b), cognitive bias relating to base rate neglect (Kahneman et al 1982) could affect individuals’ subjective assessments in the same way. Pratt’s explanation focuses on the specific information without noting the underlying base rates (the frequencies of the two groups mentioned). This base rate neglect leads him to draw inappropriate inferences regarding the relative safety of a small group (all student pilots who are post-solo but pre-qualification) versus a clearly much larger group (all instructors and professional pilots). The UK gliding movement has a single national magazine publication (Sailplane and Gliding). This bi-monthly publication contains a section of summaries for all recent accidents, including the experience level of the pilots involved (in hours). The BGA accident database from 2002 to 2006 (British Gliding Association 2007) contains 469 accidents to gliders, of which only 33 (7%) occurred to pilots with 10 hours or fewer. Hence, the magazine sections from 2002 to 2006 would have had to show approximately 13 accidents that happened to pilots with more than ten hours for every one occurring to a less experienced pilot. Base rate neglect would tend to predict that readers would perceive a low accident rate for low hours pilots simply because of the higher frequency of the more experienced group. More subtle types of communication of accident information through the gliding community (e.g. word of mouth) would be subject to the same mechanism. Hence it is possible that neglect of base rates contributes to the established misconception in both the literature and the community, as shown by the present study.

The effects of the availability heuristic would further add to this misperception. People tend to think that events are more probable if they can recall an incident of their occurrence (Botterill & Mazur 2004). Because most accidents happen to more experienced pilots, it is likely that instructors would have a close link to one or more of those accidents rather than to an accident to a low-hours pilot (e.g. they personally knew the accident victim, witnessed the accident, or heard a firsthand account). Any such things would increase the 'salience' of that accident. Salience of particular memories causes them to be more available (retrievable) and hence they tend to get disproportionately used as a cue to probability (Kahneman et al 1982). This would cause people to overestimate accident rates for pilots with experience (i.e. the majority) compared to those with little experience.

Base rate neglect and availability bias give strong explanations as to how instructors could come to perceive the situation as it is expressed in the literature, without necessarily having read or heard about it before. Some authors may have been biased in the same way but it has also been previously shown that at least part of the misconception in literature has been driven by studies that do not collect or properly account for, exposure data (Jarvis & Harris 2007b).

Since all 74 participants gave estimates above 10 hours, it means that all the qualitative data (the explanations given) relate to a misconception about relative accident levels, and hence can inform about why these misconceptions come about, or how they are justified by those that hold them. In line with general theories of attribution, explanations tended to relate to either the predisposition of the accident involved pilots (category 1: attitudes, n = 38) or the situational challenges faced by more experienced pilots (categories two and 3, challenging situation and reducing supervision, n = 26).

Most explanations tended to focus on the predispositions of accident pilots (i.e. aspects of their attitude), possibly due to literature claims of this kind and possibly due to the known tendency of people to attribute personal predispositions to events rather than situational characteristics (Gilbert 2002). In the same way that instructor estimates reflected the figures in the literature, explanations given for them are also very similar. Statements coded in the categories relating to pilot attitude used commonly recurring terms such as “over confidence”, “risk taking” (or seeking etc), “pushing limits” (or boundaries etc) and “complacency”. These are remarkably similar to those used in the literature (see Booze 1977, Telfer 1989, O’Hare 1999). Overall this shows that most gliding instructors rely on so-called ‘folk models’ (Dekker 2006) to ‘explain’ problems encountered by their relatively experienced colleagues.

It is clear that the perception of UK gliding instructors regarding the effect of experience on safety is not only inaccurate but also misguided. Jarvis and Harris (2008) showed that most accidents to experienced pilots (over 10 hours) between 2002 and 2006 occurred away from the base airfield. Chapter six showed that these were mainly involved with problems occurring during off-field landings and hence were strongly associated with more challenging situations encountered only by pilots with the necessary experience and qualifications. Hence there is a fundamental situational factor underpinning accidents to experienced glider pilots, despite most participants believing that pilot attitudes are the main issue for this group. Only a fifth of participants cited reasons related to the situational challenges found to be associated with accidents at this level (see ‘Results’ section B, category 2). The main challenges referred to were ‘cross-country flying’, ‘competitions’ and ‘field landings’. Hence these participants were accurate in that the reasons they gave were in line with their estimates, when compared to accident data. It is therefore likely that their misconception of accident rates was driven by simple frequency bias (e.g. base rate neglect), or by engrained misunderstanding, as opposed to poor reasoning.

The other category made up of situational factors consisted of explanations to the effect that accident rates went up as post-solo supervision or other restrictions were lifted. This implies that the participants believe post-solo supervision has a protective effect. The mean numerical estimates given by this group were predictably much lower than for other groups (mean = 66 hours, see Table 8.2). This is expected because solo supervision starts reducing after pilots have several solo flights and is comparatively negligible after pilots have achieved their Silver C qualification (for most this would be achieved sometime between 20 and 100 hours). Most of the participants who estimated a figure of 50 hours or less gave explanations belonging to the 'post-solo supervision' category, see Table 8.2). Since no research has been done specifically on this experience bracket, it is not known to what extent the reduction of supervision effects accidents at this level. Accident rates for early solo pilots (Jarvis & Harris 2007b) show that post-solo supervision does not compensate for the increased vulnerability of early solo pilots, in terms of accident prevention.

Flight Phase Rankings

Results from the ranking of flight phases show that for four flight phases (pre-flight, top of launch to start of circuit, circuit and landing) instructors ranked the likelihood of a causal accident event as being the same or close to that derived from the accident data on inexperienced solo pilots (Jarvis & Harris 2008). However the mean rank for the launch and approach phases suggest that instructors overestimate the relative probability of an occurrence during the launch and underestimate the occurrence during the approach when pilots first fly solo.

This could be partly explained by the previously outlined 'dread' factor as well as the availability heuristic. Table 5.11 (Chapter 5) shows that accidents coded as being caused during the launch phase were the most likely to have serious consequences (measured across all UK glider pilots regardless of experience level, and hence relevant to instructors flying solo). Despite being the second lowest accident causing flight phase (after pre-flight) it had the highest number and rate of fatalities (1 in 10). Additionally 38% of these accidents ended in injury (15 out of 39), compared to the next highest phase 'pre-flight' (33%), then 'approach' and 'general flying' (both 25%) and 'circuit' and 'landing' (both 9%). Hence for the instructor, launch events represent the causes of accidents which are the most likely to end in death or injury to themselves. Specifically the rotation into the climb on a launch is associated with most fatal accidents (Jarvis & Harris 2008). After a number of serious accidents, in 2007 the BGA launched a publicity campaign about the serious consequences of winch launch accidents and issued a leaflet for every UK glider pilot about the dangers of rotating too quickly into the climb. Hence accidents caused by winch launching have been recently refreshed in the minds of instructors and are therefore more likely to be used to judge frequency, based on the availability heuristic. In addition, the severity of consequences make these more salient, which increases the effect of the heuristic. Because these accidents are spread across all experience levels they could equally happen to instructors. This could cause worry to instructors which means that their high rankings for this phase could have been contributed to by the 'perceived dread'.

Conversely, the approach phase was not associated with any fatal accidents in the years 2002 to 2006. Interestingly however this was also true of the landing phase, which had a mean rank of five and yet had the lowest rates of injury and a lower accident rate for more experienced pilots than the approach phase (Jarvis & Harris 2008).

On the face of it this appears to be a strange result. However one possible explanation for it can be found in Table 5.10 (Chapter 5); landing was the cause of a disproportionate number of accidents during training (i.e. occurring to instructors) compared to all other flights by experienced pilots. It is also notable that the second highest number of instructing accidents was associated with the launch phase. As previously explained (chapter 5) these findings are probably because of the proximity to the ground and time to recover when students make errors in these two phases. This further helps to explain the possible feeling of 'perceived dread' experienced by the instructor when considering the launch phase, particularly considering the unforgiving nature of these accidents. It is quite possible that launching is an area where instructors feel particularly vulnerable to errors made by students.

It is possible that the landing phase also features highly in the instructor's mind for similar reasons, although with far less chance of a severe outcome. The frequency of landing accidents has been shown to be far higher than launching accidents for early solo pilots as well as instructional flights (Chapter 5). Hence the instructor possibly has to react quickly on numerous occasions to prevent landing accidents during instructional flights. This is supported by the findings outlined in Chapter six that showed misjudgement of the landing flare to be a common cause of accidents in early solo pilots. It is probable therefore this misjudgement also occurs very regularly during instructional flights and such errors require quick reactions from an instructor, and may very possibly be remembered for their salience! These effects would be weaker for errors made during approach during instructional flights, as well as for the circuit and proceeding upper air work. The height and time available would normally allow the instructor to deal easily with errors or even let the student correct their own errors. All this would predict that instructors are able to easily retrieve from memory many instances of landing errors made on instructional flights, which helps to explain why they rank landing as the second highest phase but approach as the second lowest.

All of this indicates that instructors may not base their evaluations on an objective assessment of what students would do when solo, but on the factors that potentially have most negative impact upon themselves, either while instructing, or flying generally. This is similar to findings by Bohm and Harris in which the observations of machine operators were more concerned with 'perceived dread' than risk per se (Bohm & Harris 2009a). Whereas the high rankings given by instructors to the launch and landing phases are understandable, the low ranking of the approach phase is of concern, particularly considering that this phase is associated with a disproportionally high rate of accidents for early-solo pilots (Jarvis & Harris 2008).

Overall

These results demonstrate that gliding instructors do not realise that the most likely pilots to be involved in accidents are those that they have just sent solo.

This should raise serious concerns about the basis of decisions made by instructors to send student pilots solo. Such 'solo decisions' could be questioned from the point of view of the accident rates alone (Jarvis & Harris 2007b) but the present study provides strong evidence that as well as underestimating the probability of accidents occurring during such flights (relative to other flights), instructors may not give sufficient consideration to the most problematic areas of flight for early solo pilots (specifically the approach phase). Their assessments of the competence of newly soloed pilots may be strongly affected by biases such as availability bias and 'perceived dread' as applied to their own flying. Since it has been shown that instructors cannot easily judge the areas of danger faced by early solo pilots (in terms of flight phases) it opens up the question of what instructors are basing their decisions to allow solo flight on.

It is particularly curious that the approach phase was not ranked as a relatively more important area by instructors, considering its disproportionate accident rate (Jarvis & Harris 2008). Chapter six showed that a relatively common accident scenario for newly soloed pilots was insufficient reduction of airbrakes (e.g. leaving them deployed while a serious undershoot develops to the point where the glider can no longer reach the intended landing area). In line with the pyramid theory of accidents and incidents (Heinrich et al 1980) if high numbers of gliding accidents are caused by certain events then many more such events probably occur that do not lead to accidents. Therefore it is likely that the same sorts of errors and events causing accidents for early solo pilots also occur in the lead up to solo (i.e. during training). The question therefore is; do instructors notice them, and do they take them into account in their assessment of student performance? It could be argued that if instructors were noticing a high number of such 'undershoot' occurrences then they would have ranked the approach as a more critical area. Interestingly the approach phase did not feature highly in instructional accidents between 2002 and 2006 (see chapter 5) and no instructional accidents were caused by insufficient reduction of airbrake (see chapter 6). This could be because instructor intervention or prompting means that such errors do not become an issue during instructional flights. As previously stated, it could be that some errors made on instructional flights are more salient than others, since some are more recoverable. This would have implications for the instructor's assessment of a student's flying and the decision to allow solo flight in the first place.

Research is required into which aspects of student performance instructors are basing solo decisions on. This might help to address the very high accident rate of early solo pilots. The following chapter will look more deeply into this critical decision process.

Chapter 9 - Investigation into the factors influencing instructor decisions to allow and disallow solo flight

Introduction

It was found that gliding instructors under-estimate the relative likelihood of accidents faced by early solo pilots (defined as those with 10 hours experience or fewer) (see chapter 8). Out of a sample of 74 instructors, none realised that low-hours pilots were the most likely glider pilots to be involved in accidents, as shown by Jarvis and Harris (2007b). Instructors incorrectly believe that pilots become more likely to have accidents as they gain experience up to a point between about 200 and 500 hours (see chapter 8). These consistent beliefs are based on notions of unsafe pilot attitudes, increased challenge and reduction of post solo supervision, and they reflect similar claims in aviation literature. It was proposed that several cognitive biases could contribute to these misunderstandings. Being as the most vulnerable group of glider pilots (statistically) are those who have just been sent solo, it is of concern that those sending them solo may underestimate the likelihood of an accident.

Additionally, instructors' perceptions of the areas that are most likely to cause these accidents (in terms of flight phases) are not aligned with the accident data. Specifically, despite the approach phase being the most likely to cause accidents to newly soloed pilots (Jarvis & Harris 2008) instructors perceive it as the least likely to do so relative to all other in-flight phases.

The results of the previous study (chapter 8) raised potential concerns with regards to instructor judgment in terms of sending student pilots solo. It was suggested that their decisions may be strongly affected by biases such as availability bias and 'perceived dread' in the same way that their perceptions of accident likelihood appeared to be. Simply put, since instructors appear to wrongly perceive the main issues faced by early solo pilots (in general terms as well as by flight phases) it opens up the question of what instructors are basing their decisions on to allow solo flight.

Chapter six showed that accidents to early-solo glider pilots (under 10 hours) were confined to just a few causes. These were misjudgements of the approach path (leading to undershoots), misjudging the landing round-out, and control of airspeed and pitch on approach. A well known rule of accidents and incidents (the pyramid theory), based on extensive accident research, is that for every accident there are roughly 30 occurrences of similar unsafe practices or conditions (Heinrich et al 1980). This theory would suggest that the aforementioned accident causal events should occur regularly in the period leading up to solo flying. If so then one would expect these factors to be present in training accidents, but the picture does not fully support this.

Research literature looking at accidents occurring to general aviation pilots under instruction gives support to the general point; events on landing and during 'touch-and-goes' were associated with the highest frequencies of accidents in both instruction and early solo flying (Baker et al 1996). This is similar to the findings of Jarvis and Harris (2008) that in UK gliding landing was the most frequent flight phase for seminal events for both early-solo pilots and student pilots under instruction. However, further study by Jarvis and Harris (chapter 6) showed that the landing events on instructional flights were only associated with misjudgement of the landing flare, and specifically leaving it too late, whereas for early-solo pilots flaring too early was just as likely to cause an accident. Accidents initiated during the approach phase were not highly associated with instructional accidents, whereas they were associated with the highest rates and totals of seminal accident events for early solo pilots.

Possible answers to these apparent anomalies lie in the contribution of the instructor. Certain errors made by students during training might be less likely to lead to accidents than others (due to the instructor contribution) meaning that these errors never enter the accident data and so are invisible to these types of studies. This theory is supported by the finding that of the specific factors that caused accidents to early-solo pilots, the only one to feature highly in training accidents was 'initiating a late landing flare'. Clearly there is very little time for an instructor to prevent an accident when the seminal event occurs close to the ground in these circumstances; "the worse problem for both the student and the instructor occurs when the initial round out is left too late" (Piggott

1991). This helps to explain the high number of these sorts of accidents occurring during instructional flights. The other types of accidents occurring to the early-solo pilots (overuse of airbrakes / mishandling on approach, and landing flare too early, chapter 6) are not represented at all in the instructional accident data. Notably these events all occur at a higher altitude during the flight (than a late landing flare) allowing the instructor time to take over control to prevent an accident. These particular kinds of errors are therefore less likely to lead to accidents with an instructor on board, than late landing flare events. In this way such errors may never appear in the training accident data or subsequent analysis. Hence they only appear in the data when pilots are flying solo despite possibly occurring right through their training.

Despite some apparent anomalies in accident cause between student pilots and solo pilots, there is a strong case (supported by the nature of those same anomalies) to explain how only those that leave the instructor less time to respond appear in the accident statistics for student pilots.

If the causes of accidents occurring to low-hours pilots were more evenly distributed among flight phases and causal categories, then it would raise questions about the general standards of new solo pilots. However, as shown the accidents are not spread in this way, but are clustered around a small number of relatively discrete areas which, with increased pilot experience become relatively insignificant in terms of accident numbers. This raises the question of how much weight is put on these factors by instructors when judging student pilots prior to sending them solo or whether these areas are considered at all. Research is required into how instructors make 'solo-decisions' and what aspects of student performance they base these decisions on, in order to address the very high accident rate of early solo pilots.

There are no standards laid down as to what constitutes a solo-standard, nor what instructors should look for. The British Gliding Association solo training syllabus; 'BGA Solo Certificate' (another name commonly used for 'the A certificate') provides guidance on what exercises should be completed by students before the solo certificate is issued. The list of training exercises is given as 'lookout', 'effects of controls', 'use of trim', 'the straight glide', turning', 'airbrakes', 'approach control', 'landing', 'circuit planning', 'launching', 'stalling', 'spinning and spiral dives', and 'first solo' (British Gliding Association 2008b). The latter is simply described as "Normal take-off, circuit and landing". Instructors are referred back to the BGA instructors' manual for detail on all these exercises.

The requirements to fly solo simply state that a candidate must be 16, have signed a declaration of medical fitness, and must have completed 20 flights (British Gliding Association 2008a). The rules state that tests (leading to certificates) "can only be done when the candidate is alone in the glider" (British Gliding Association 2008a). The first test is to obtain a certificate of competence termed the 'A' badge, which consists of "one solo circuit… followed by a satisfactory landing" (British Gliding Association 2008a). It is difficult to imagine what constitutes a failure to this test other than an accident, which should bring its value into question. Hence the solo standard appears to be recognized retrospectively, in that a pilot is deemed competent to fly a solo only after they have flown solo! The reason for this rather odd requirement appears to be historic, and comes from a time when there were no two seat gliders, and so the first flight was also the first solo, and hence no 'solo-standard' existed. Strangely the assessments have remained the same for over seventy years, and the system of testing illustrates the lack of attention to pre-solo achievement. The findings from chapter eight along with claims by popular textbooks, suggest that the UK gliding movement (from a training perspective) does not perceive a safety issue with newly soloed pilots. This would help to explain the lack of attention given to the issue of determining whether or not pilots are ready for solo.

The BGA instructors' manual includes only a short section devoted to the first solo, but there is no mention of the decision itself apart from the following: "At first sight the decision to send a trainee on a first solo looks to be a tricky one. It isn't, because if you aren't sure you don't do it" (British Gliding Association 2003). This advice assumes that gliding instructors can make reasonable assessments of the risk (i.e. to be 'sure') but it has been suggested that this is not true (see chapter 8). However the BGA syllabus and instructors' manual (British Gliding Association 2003) do include performance markers to be attained before solo. These include circuit ‘planning’, approach control, spot landings, stalling, spinning, speed and directional control, launch failures, satisfactory take-off and launch and thermal soaring. Each of these has a very short description about what is required. Of particular interest are approach control and ‘spot’ landing, which represent the two areas found to cause most problems in early-hours pilots (Jarvis & Harris 2008). The former is accompanied by the description “able to recognize and correct for an undershoot, and understand the need to turn in early if necessary” (British Gliding Association 2003). This description does appear to be in line with accident trends showing that judgment is problematic on approach, leading to undershoots caused by overuse of airbrakes. Although there is no further advice on how instructors should assess the ability of the student to recognize and correct for undershoots, it indicates that instructors should look at this aspect of a student’s flying when making a decision about their standard prior to solo. In terms of landing, it is interesting that ‘spot landing’ is given instead of just ‘landing’, and the only advice is that this is “not necessary, but should be done competently if the glider has good airbrakes” (British Gliding Association 2003). Landing itself, including judgment of the flare is not included in this list, although ‘ballooned landings’ are identified as one ‘common difficulty’ experienced on first solos, in another section. Considering the rate of flare misjudgments leading to accidents this could be a serious omission. The instructors’ manual deems that the only other areas that trainees regularly have difficulty with on first solos are ‘getting the circuit too close’, and ‘making sloppy turns’ (British Gliding Association 2003). There is no mention of the approach phase in this context, which again appears to be a serious omission given the findings of Jarvis and Harris (2008), and those outlined in chapter six.

Approach control is covered in a separate section of the manual and includes a large variety of exercises for instructors to do in order to show problems to students that can occur (including undershooting and the effect of not closing the brakes to recover from it). These are written as instructor demonstrations however (rather than student exercises), and the only exercise that is about checking the student for undershoot tendencies involves the instructor setting up a low final turn to check that the trainee does not open the brakes immediately (termed ‘landing lever syndrome’). The BGA instructor training syllabus includes six approach control exercises but these are also instructor demonstrations, including ‘undershoot and recovery’ and ‘progressive undershoot’ (British Gliding Association 2008c). The instructor must learn to do these competently in order to demonstrate them to student pilots. There are no such exercises for the diagnosis of issues in pre-solo pilots, for any specific area of the approach. Indeed there are only two exercises out of 55 on the BGA instructor training syllabus (Appendix G) that are concerned with diagnosis or observation of a student’s flying as opposed to the flying skills of the trainee instructor. These two exercises come under the heading of ‘fault finding’, and are termed ‘handling skills’, and ‘judgement exercises’. No more information is given on what these might entail. The syllabus contains nothing about the decision making process of sending a pilot solo.

Gliding and general aviation text books (written for pilots) include even less information useful to instructors, or students wishing to attain the standard. Most simply include encouragement to the solo pilot such as “When your instructor steps out of the aeroplane and leaves you to your first solo flight, you are being paid a big compliment” (Thom 1997), and “Relax as much as possible and let your training take over. Enjoy every moment of this flight - there is only ever one first solo” (Pratt 1994). Additionally all these books include phrases that reassure a student pilot about the instructors ability to judge when a pilot is ready to fly solo, for example “remember that he is trained to judge the right moment to send you solo” (Thom 1997) and “Your instructor will not send you first solo unless he or she is totally confident of your ability” (Pooley 2003). However there is no description or advice on what this standard actually is, or how such confidence or judgment is attained by the instructor. The most that Thom (1997) states on this point is that “The usual standards apply to your take off, circuit and landing… your instructor, when sending you solo, not only considers you competent to fly a circuit with a normal take off and landing, but also considers you competent to handle an emergency”. Pooley (2003) also mentions emergencies in a statement that the student would have perfected their landing technique and practiced various emergency procedures, including engine failures after take-off and go-around. Pratt (1994) offers little more by stating that “Your instructor will send you solo based on his judgement of the safety and consistency of your flying and judgement”. The findings from the previous study (Chapter 8) strongly suggest that instructors' 'judgement of safety' as applied to this pilot group is poor, which must bring into question the confidence of the above claims.

Gliding texts tend to elaborate on the personal attributes needed to fly solo more than the human factors or technical skills. Piggott (1997) states that a solo pilot should possess 'confidence about his ability and be aware of his limitations', as well as being 'experienced and confident in his ability to cope with emergency situations' and have 'an honest and mature outlook'. He only states a few technical skills; 'prove that you can plan your circuits and deal with any contingencies such as cable breaks, running short of height, and stalling and spinning, without advise or help from the instructor" (Piggott 1997). Hence, ability to deal with emergencies appears to be the most consistent theme running through training literature, as to what solo competence looks like. No texts for pilot training go into depth on any of these issues.

It is clear that there is a serious lack of guidance in training literature and texts on how to decide when a student pilot is ready to fly solo in general, as well regarding specific areas of the flight.

Unfortunately there is almost no research into this area within the scientific domain either. The only study to look at instructors decisions regarding solo flight was in general aviation, and required twenty instructors to create decision 'policies' of what a student should attain to be competent for solo flight, based on their scores on the 16 elements in the training syllabus (Ikomi & Guion 2000). Consistencies were found in terms of instructors ruling out solo flight unless a score of at least three out of five had been achieved (average) in virtually every maneuver in the syllabus, which suggested a cut-off point being set on each item. Although it was noted that 'policies' were very different, even within a single flight school, it was concluded that more flying instructors showed insight into their decision making than football coaches or psychology professors, based on matches between verbal reports and actual judgements (Ikomi & Guion 2000). However the numbers were not significant and there was also extreme variability in the willingness of instructors to send students solo.

The limited conclusions of this work could be partly due to research limitations. There are questionable aspects to such a study. It begins from the assumption that instructors base their decisions of whether to allow a student to fly solo or not, on those elements found in the training record (i.e. that the decision 'attributes' are all contained on the record sheet). The decision process involved in sending a pilot solo could be more complex than this, and involve other less obvious decision dimensions and non-declarative knowledge. The use of non-declarative knowledge in decision making is sometimes referred to as 'intuition', which Klein (1998) describes as "the use of experience to recognise key patterns that indicate the dynamics of the situation". Expert knowledge is stored in the 'knowledge-base' (long-term memory) which is an unconscious resource with a vast array of specialised processors called schemata, used to process familiar information rapidly (Reason 1990). Schemata are activated by triggers (often unconsciously) and the more that a schema is used, the less is needed to trigger it "particularly in very familiar environments" (Reason 1990). Hence instructors, familiar with the environment, may rely on such unconscious cues to judge when a students' flying is not 'typical' or 'satisfactory' in order to help make an assessment as to whether or not the student is safe to fly solo. Intuitively judging whether a situation is 'typical' in this way is a process that is often relied upon in expert decision making tasks (Klein 1998).

The work of Ikomi and Guion (2000) may therefore have missed some important decision attributes by asking instructors to base their 'solo decisions' only on the scores of the elements on the training record (on paper). All instructors claimed to have used non-compensatory combining rules when making decisions based on these elements (Ikomi & Guion 2000). Strategies of decision making are often characterized as either compensatory or non-compensatory (Rothrock & Yin 2008). Compensatory decision making theories postulate that people make numerous and exhaustive 'trade-offs' between attribute cues (Rothrock & Yin 2008), and as such are based on the classical model of rational choice. According to this model, the "rational" actor chooses the best options to follow by assessing the probability of each, then discerning the utility of each one, and combining these two assessments (Gilovich & Griffin 2002). However this model has regularly been shown to be unrealistic, particularly in terms of complex judgements. The theory of 'bounded rationality' (Simon 1957) led to a large and ongoing body of work concerned with how people short-cut the rational decision process, of which non-compensatory theories are an important part.

Simon proposed that instead of valuing and combining decision attributes fully, people simply set 'aspirational levels' on various dimensions, meaning thresholds with which they would be satisfied. This process is known as 'satisficing', and reduces the effort and time required by the process. In the context of the 'solo decision', satisficing behaviour would relate to the instructor 'setting' a level that they will be satisfied with on any particular attribute (e.g. the landing) rather than fully evaluating it. If this were the case then there would be little value (simply in terms of the 'solo decision' outcome) in a student making a very good landing as opposed to a satisfactory landing, since either would satisfy the 'landing attribute' part of the decision process.

Non-compensatory combining rules mean that one decision attribute cannot be compensated for (in terms of the decision outcome) by another. For the research of Ikomi and Guion (2000), the non-compensatory nature of the decision meant that unsatisfactory performance on one part of the flight training record (less than three out of five) could not be compensated for by good performance on another. Therefore the score of three represented a 'cut-off' threshold of satisfaction for each manoeuvre, in line with satisficing behaviour within a non-compensatory model. A score below three on any manoeuvre could not be compensated by a score above three on another.

Hence the use of non-compensatory models in 'solo decisions' would appear very possible because any element of a flight could cause an accident, and one satisfactory element (e.g. take off) could not compensate for an unsatisfactory one (e.g. Landing). Hence the rule proposed by Ikomi and Guion (2000) is theoretically applicable, and the cut-off points (scores of three on each element of the record card) would equate to instructor satisfaction and hence suggest the possibility of satisficing behaviour.

A second interesting feature of the research by Guion and Ikomi (2000) was that a successful first solo was used as a definition of a correct judgement on the part of the instructor to allow solo flight. The implications of allowing a pilot to fly solo are arguably far reaching, and may lead to accidents several flights down the line (for example when the pilot experiences their first emergency, or first challenging weather situation). This is supported by the findings that pilots with up to 10 hours have twice as many accidents per launch than their more experienced counterparts (Jarvis & Harris 2007). Therefore an uneventful first solo flight is arguably not in itself an indicator that a pilot was in fact ready or competent to fly on their own, and therefore not necessarily a good singular indicator of correct judgement on the part of the instructor.

Considering that the approach has been associated with most accident causes in low-hours pilots (Jarvis & Harris (2008), it of note that the training records used in the research of Ikomi and Guion (2000) did not include the final approach as an element to be scored and taken into account in the solo decision. It may be that the approach phase was included with the landing phase, but that is highly unlikely based on the description of that category on the record used. Additionally there was no element for the circuit, and so it appears that the final approach was not judged at all by the instructors as part of the research, even though it has been associated with a number of accidents in general aviation (Baker et al 1996). This would appear to be a serious limitation caused by the research design and provides a strong case for a research methodology that does not assume the decision attributes used by instructors. Since it is a relatively unknown decision process, an exploratory technique is required in order to establish the attributes of the decision rather than imposing them on the participants. This will provide a foundation of knowledge for later work that could begin to map the dynamics and measure the weightings of the decision process.

The decision to allow solo flight is a highly critical one, with possible fatal consequences if made incorrectly. Although many consequences will show up early in a pilots career (during the first ten hours of solo flight, as shown by previous findings), it is also possible that lack of skills could lay dormant until a situation arises where incorrect technique combines with circumstances to produce fatal results much further downstream. Therefore instructors have a huge responsibility to correctly judge whether a pilot has mastered all the necessary skills to be safe, and so one would expect a reticence in regard to their decisions to send pilots solo. Indeed, the simple need to be absolutely sure is one of the few pieces of guidance in the training literature as previously outlined (British Gliding Association 2003, Pooley 2003). Interestingly the results of Ikomi and Guion (2000) show an unexpected leniency on the part of flight instructors in sending pilots solo, although this may be a reflection of the validity of the research method.

Given the lack of research or knowledge about this complex decision, an exploratory methodology is required that can establish how instructors decide that a pilot can fly alone. This may provide some answers with regard to the very high accident rate of early solo glider pilots.

In order to inform any remedial actions to the high accident rate of early solo pilots, exploratory research is needed that can build a picture of what aspects of students' performance cause instructors to allow them or prevent them from going solo, and how that picture overlays the pattern of accidents occurring to these pilots after the decision has been made. No research has been done that answers these questions.

Aims and Objectives

The first objective was to discover what elements gliding instructors identify as being satisfactory and unsatisfactory performance in their students, and establish how these factor into the decision to allow or disallow solo flight. The second objective was to compare such findings to previous findings from accident data (in terms of areas of concern).

The overall aim of the study was to gain an initial understanding of instructor decisions with regard to solo flight, and to explore in what ways instructor decisions could be contributing to the high accident rate of newly soloed glider pilots.

Method

Pilot study

Since nothing was known about the process of instructor decision making regarding solo flight, an initial exploratory pilot study was undertaken. Twelve open interviews (simply asking instructors to talk about the decision) revealed that they felt it was a hugely complicated process with many factors. They found it difficult to explain the whole process, and in many cases stated that they simply did not know how they came to a conclusion. A large number expressed the idea that while they found difficulty in explaining reasons behind the decision to send a pilot solo, they would find it easier explaining why they decided not to send pilots solo. A number of instructors talked clearly about observable aspects of student flying that had led them to send pilots solo, or not allow solo flight. Most instructors said that the decision was largely based on their observations of the student's flying, but when asked to elaborate on the detail they often ran into difficulty. The interviews also revealed numerous possible decision dimensions, some general and some specific (related to aspects of the students' performance).

Research Questions

From the pilot study findings, two broad research questions were formulated that were exploratory in terms of finding out the basis upon which instructors decide to allow student pilots to fly solo, and whether these aligned with the issues raised by the accident data.

1. What factors do gliding instructors identify as being satisfactory and unsatisfactory performance in students, in terms of allowing them to fly solo?
2. How do these findings compare to previous findings from accident data (in terms of areas of concern)?

Methodological considerations

A research methodology was required that would enable open exploration regarding the decision to allow solo flight and generate data capable of meaningful comparison with previous results. Based on the pilot interviews, the Critical Incident Technique (CIT, Flanagan 1954) was chosen. CIT was ideal for probing the factors impacting upon instructor decisions without needing to discuss the decision making process directly, which pilot study participants had found difficult. Additionally, CIT allowed collection of data comparable to previous accident analysis.

Critical Incident Technique (CIT)

CIT was developed by John Flanagan in the 1940s as part of the Aviation Psychology Program of the US Army Air Force (USAAF) and is widely associated with studies of human error (Kirwin &Ainsworth 1992). The technique is firmly established and has been extensively used over the past 60 years (Butterfield et al 2005). It is recognized as an effective technique when researching a topic that is little understood or sparingly documented Gremler (2004). Research has demonstrated that material collected using the method is highly representative of the essential points required for a task, as determined from other sources (Andersson & Nilsson 1964). It is now generally accepted that CIT has proved both reliable and valid as a method of generating a comprehensive description of a content domain (Ronan & Latham 1974, Woolsey 1986).

CIT does not consist of a single rigid set of rules, but is a flexible set of principles to be modified to the task in hand (Flanagan 1954). The technique generally involves collecting first-hand reports of satisfactory and unsatisfactory execution of an assigned task (Flanagan 1954). 'Critical Incidents' themselves are effectively anecdotes describing: "behaviour either outstandingly effective or ineffective with respect to attaining the general aims of the activity" (Flanagan 1954). Thus this method was ideally suited to the present research question since there was a clear criterion for

successful and unsuccessful task execution in the form of the student's performance being judged to be good enough or not good enough to allow solo flight.
Flanagan (1954) described five important stages to performing the CIT which have been utilized in many hundreds of studies since.

- **STAGE 1 - The General Aim:** This is a functional description of the activity and specification of what it is necessary to do and not to do in order to achieve successful or effective participation (Flanagan 1954).
- **STAGE 2 - Plans and Specifications:** This stage involves "a delimitation of the situations to be observed", including information about place, persons, conditions and activities (**Flanagan 1954).**
- **STAGE 3 - Collect the data:** Data collected in a critical incident study should provide complete coverage of the content of the domain (Woolsey 1986). This usually involves eliciting critical incidents from participants in the form of interviews or questionnaires.
- **STAGE 4 - Analyse the data**: Analysing CIT data involves "analysis of thematic content, arrived at by inductive reasoning" (Woolsey 1986). Both qualitative and quantitative methods have been used in critical incident studies.
- **STAGE 5 - Interpret and report findings:** This is involves reporting findings (results) and determining the meaning of these within the context of the subject area under investigation.

All these stages were taken into account in order to complete the procedure. They will be discussed further after the fundamental procedure has been outlined.

Sampling considerations: Participants

Recent critical incident studies reveal a wide range of sample sizes in terms of participant numbers. Most modern studies use between 25 and 75 participants (see Zaidman-Zait 2007, Borgen & Maglio 2007, Coleman 2006, Neupert et al 2005). However this is not universal; Butterfield and Borgan (2005) used only 15 participants whereas Garn and Cothran (2006) used 191.

Key CIT literature generally proposes that the number of critical incidents rather than the number of participants is the key factor in sampling (Butterfield et al 2005, Woolsey 1986, Flanagan 1954) and generally little attention is given to participant sampling numbers in the research literature. This is surprising given that participant characteristics determine to whom the study can be generalised (Woolsey 1986). Hence, in order that the results could be confidently generalised across UK gliding instructors, consideration was given to obtaining a representative sample of participants.

Sampling considerations: Critical Incidents

In the same way that participant sample sizes are not uniform across studies, there is no set rule for the number of incidents required (Butterfield et al 2005, Flanagan 1954). The number of incidents depends on the nature of the task, and could range from fifty up to several thousand for particularly complex tasks (Flanagan 1954). A general rule is that incidents continue to be collected until redundancy appears (Woolsey 1986). A running count of critical behaviours emerging from the data is often advocated to find out when the addition of incidents no longer leads to the emergence of more than a few critical behaviours (Woolsey 1986, Andersson & Nilsson 1964, Flanagan 1954). An often quoted rule is that when 100 incidents only lead to the emergence of two or three new categories then the topic has been sufficiently covered (Flangan 1954). Because this was conducted concurrently with data analysis, the process is covered in STEP C (Data analysis).

Data collection considerations

The retrospective face-to-face interview remains popular for CIT studies across many domains (Borgen & Maglio 2007, Dollarhide et al 2007, Zaidman-Zait 2007). Other methods such mailed questionnaires (Garn & Cothran 2006, Coleman 2006), telephone interviews (Oldenburger et al 2008, Cottrell et al 2002) and on line survey forms (Papadakis 2008) have also been used, and there is evidence that some of these methods give satisfactory responses (Converse et al 2008). Even so the interview aligns well with the CIT method’s requirements to understand the implications of the event within its full context (Schluter et al 2007). As well as allowing researchers to follow up on interesting responses (Robson 2002), interviewing facilitates the active support of the interviewee, and allows a multi-dimensional picture to emerge, including the use of body language (Stitt-Gohdes et al 2000). Given the subject matter, visual gestures were likely (e.g. to imitate the path of the glider) and face-to-face interviewing would allow these to be noted on the transcript if necessary. Although focus groups have been used for CIT studies (e.g. Keatinge 2002), this might have inhibited responses in light of the subject area under discussion, and would not allow confidentiality.

A review of CIT studies since 1987 found that virtually all of them used retrospective reporting, as opposed to direct observation of incidents (Butterfield et al 2005). The criterion for accuracy is based on the quality of specific incidents recounted in terms of fullness, clarity and detail (Butterfield et al 2005) and hence the accuracy of participants' memory and recall is important. Where observers report on others (such as is proposed) recency is important in order to maintain accuracy (Woolsey 1986). This problem can be minimizing by restricting reported incidents to observations within the previous six to 12 months (Ronan & Latham 1974). Mitigation to this is to give participants pre-warning of the questions (Schluter et al 2007) which have proved effective. However, gliding clubs are social places where instructors might have discussed incidents between themselves and hence introduce bias (from a cognitive recall and social perspective). It was therefore decided as unwise to give much warning. Whilst some participants did clearly struggle to recount incidents during the pilot interviews, most were able to recall several in detail.

In terms of the interview questions themselves, wording is known to be crucial, and even small differences can seriously affect the range and quality of reports (Flanagan 1954). For this reason the question (and most particularly the CIT part) were carefully thought out to be reflective of the general aim, discussed with subject matter experts, and piloted on a small number of gliding instructors to assess their focus on in relation to the general aim.

Data analysis considerations

It is normal to extract (code) critical incidents from interview transcripts. However deciding what is and what is not a critical incident (in order to extract them) is far from straightforward. Flanagan (1954) describes an incident as “any behaviour either outstandingly effective or ineffective with respect to attaining the general aims of the activity”. Other studies have described incidents simply as ‘units of behaviour’ (Andersson & Nilsson 1964).

Once extracted, the analysis of critical incidents involves “analysis of thematic content, arrived at by inductive reasoning” (Woolsey 1986). Whereas most studies use some form of qualitative analysis in this way, data collected using the critical incident technique can be analysed both quantitatively and qualitatively and both have been frequently used (Chell & Pittaway 1998).

The general purpose of analyzing critical incident interviews is to understand the commonalties among responses (Stitt-Gohdes 2000). The causal taxonomy formed from UK gliding accident data (see chapter 6) could have formed a foundation for such analysis, but this would assume that explanations of the ‘solo decision’ factors (CIT data) shared a common basis with the accident data. To make such an assumption would be premature in view of the lack of previous research in this area. Without an adequate template, the data analysis required a more fundamental emergent technique. Critical incident data can be (and has been) categorized according to the principles of grounded theory (Gremler 2004), and this methodology was suitable for the analysis of meaning within the data. In Flanagan’s original paper on CIT it was stated that “the aim of analysis should be to increase the usefulness of the data while sacrificing as little as possible of their comprehensiveness, specificity, and validity” (Flanagan 1954). To this end it was very important that the categories accurately reflected exactly what had been said and meant by the participants, and that categories were not over-generalised.

Reliability and validity considerations in CIT research

Critical incident technique has been established as a valid and reliable technique for describing content domain (Andersson & Nilsson 1964, Ronan & Latham 1974). However, as CIT has evolved it has become increasingly common to include measures aimed at convincing readers of the credibility of results (Butterfield et al 2005). Table 9.1 summarises eight such procedures identified as being used in CIT research, in the order described by Butterfield et al (2005). Although it is unusual for studies to use all such procedures, it was decided that within reasonably practical restraints all should be attempted to insure the most robust data collection and analysis possible.

Table 9.1. The eight validity and reliability processes identified by Butterfield et al (2005) as being common in CIT research.

Independent incident extraction check: A second person independently extracting critical incidents from transcripts (typically 25% of the total incidents gathered)
Participant cross-checking interview: A number of participants are interviewed again after the first round of categorisation, to confirm that the categorisation makes sense to them, and reflects their experiences. They are also asked to review their own incidents, and add, delete, or amend as they wish.
Categorisation by an independent judge: Typically 25% of incidents are placed into existing categories, after the initial categorisation is complete.
Tracking the redundancy of incidents during data collection: Batches of critical incidents are checked during data collection for newly emerging themes. When no new themes (or very few) emerge from a batch the subject area can be said to be covered.
General SME review: Submission of tentative categories to two or more SMEs in the field, to review, indicate any categories that surprise them, or are missing.
Participation rates: The number of participants citing a particular incident is divided by the total number of participants. A participant rate of 25% is suggested as valid.
Descriptive validity. This is concerned with the accurate interpretation of incident descriptions. Recording and or transcribing helps to maintain this, but also participants can be asked to check incidents within the categories to insure that the category descriptions represent what they were trying to express.
Interview fidelity: An experienced researcher listens to a number of interviews to insure that consistency is being maintained, rigour is being upheld, and checks for leading questions.

Procedure

Outline

Face to face interviews were held with a wide sample of qualified gliding instructors from a number of UK clubs with the aim of establishing the reasons that they have sent pilots solo, or decided not to. The following sequence of questioning was asked many times of each participant, and probe questions were used if more information was required.

"Think of a student that you have flown with recently under these circumstances [assessing for solo] *but don't tell me their name. Did they satisfy you that they should be allowed to fly solo that day? What did the student do or not do that made you say [no / yes] on that occasion?"*

All interviews were recorded and transcribed. The critical incidents were 'extracted' from the transcripts. In essence, a critical incident was any statement made by the instructor that gave a reason for sending a student solo, or deciding not to.
Two independent sets of analyses were then carried out on the entire set of critical incidents.

1. **Meaning analysis:** In common with previous studies, techniques from grounded theory (Strauss & Corbin 1967) along with the 'paired comparison' technique (described by Partington 2002) were used to group the critical incidents into categories based on their meaning in relation to the CIT general aim (why a student was sent solo, or nor).

2. **Flight Phase analysis:** The flight phase template from Jarvis and Harris (2008) was used to carry out template analysis (as described by King 1998) on the critical incidents. This was done in order to find out about the relationship between flight phases and 'solo decisions', i.e. how often each flight phase was related to the events driving the instructors' decision.

Throughout the process, procedures were carried out to maintain and check the reliability and validity of the process. These are shown in Table 9.1.
Finally, in order to compare the CIT ‘meaning’ analysis directly with the previous research, all events found to cause accidents in UK gliding (chapter 6) were coded into the ‘CIT’ top-level categories.

The general CIT aim

The purpose of establishing a general aim is to gain an understanding of what a person who engages in the activity is expected to accomplish (Butterfield et al 2005). This can be obtained by asking participants and should result in a functional description of the activity that is at least acceptable to potential users (Flanagan 1954).

In terms of the ‘general aim’ the task was piloting a glider (for a student pilot). However the research question did not ask about successful performance on this task, but what the instructor perceived as indicating it. This simplified the general CIT aim because the success criteria did not need to be defined due to it being part of the question.

After discussion with subject matter experts as recommended by Flanagan (1954) the general aim of the activity was established as:

> “For the student to satisfy his/her instructor that they have reached an overall standard of flying that would make it acceptable (for the instructor) to allow them to fly solo. This should be achieved through demonstrating that they can safely operate a glider in the prevailing conditions and under any restrictions laid down or accepted as part of early solo flying”.

This was shown to other subject matter experts (SMEs) and read to participants who were also SMEs by nature of being experienced glider pilots, and was found to be acceptable to all. It is important that collected reports of observed behaviour are relevant to the general aim in terms of making a significant’ contribution, either positively or negatively, to the general aim of the activity (Flanagan 1954).

Plans and Specifications

In the present study the situation was any instructional glider flight where the instructor was considering allowing the student to fly solo after that flight or session. This could be a formal flight or test (often termed a check flight, or check 'ride') or any instructional flight leading to the consideration. External factors such as weather conditions, traffic density, logbook comments and progress cards might impact on the instructor's decision outcome but it could assumed that if the instructor was genuinely considering whether to allow solo flight then these must have been satisfactory. Hence the observed behaviours of interest would be only those attributable to the student during that session, and assuming the instructor was already reasonably satisfied about the student's progress and background. An explanation of this was included in the interview schedule, along with specific information to guide participants so that they knew exactly what was required of them, and within what context.

Sample

Emails were sent to instructor populations of four UK clubs and several visits were made to each. Instructors were then interviewed as was practical on location (almost all UK instructors are volunteers, and it was not always possible to know exactly who would be available and when). This was successful because the eventual sample showed a large spread of instructors representing instructional activity at eight UK gliding clubs, including flat sites, coastal sites and ridge sites.

Interviews

For the reasons given previously; individual face to face interviews were used, containing retrospective, open-ended questions. Participants were urged to try to use incidents from the previous 12 months. Where possible they were given a small amount of notice of the questions (no more than a day) which gave them a small amount of time for recalling incidents prior to the interview.

The full interview schedule as shown in Appendix H was finalised after a small pilot study of four participants using an initial schedule. This pilot interview process is outlined in Appendix I.

A three stage questioning process was found to work best. The process worked as follows.

1. An instruction was given (after the pre-brief); “Think of a student that you have flown with recently under these circumstances, but don’t tell me their name.”
2. Only after acknowledgement from the participant, a further (secondary) question was then asked; “Did they satisfy you that they should be allowed to fly solo that day?”
3. The real CIT question was asked; “What did the student do or not do that made you say [no / yes] on that occasion?”

In trials, it became evident from the responses to the yes / no question (part two) that a number of instructors had not considered a specific case as requested. The subsequent (yes/no) question was therefore important because it could be used to help maintain the interviewees focus on that specific case.

Interview probes (see Appendix H) were included based on Rous and McCormack (2006), as well as probes aimed at discovering what part of the flight the incident occurred in (if unclear).

All interviews were digitally recorded and most parts faithfully transcribed. Although transcription is not essential in CIT (Woolsey 1986) this did assist with maintaining descriptive validity (see Table 9.1).

Data Analysis

In common with most CIT studies, data analysis was conducted concurrently with data collection in order to establish the scope of the task and determine the number of incidents required. Before data analysis could take place, the critical incidents had to be extracted.

Incident extraction - development of the method of extraction

A method of critical incident extraction was developed by two independent raters jointly coding several transcripts. The initial brief was deliberately loose, and centred around the CIT ‘general aim’:

> “Any statement or part of a narrative that was deemed to be meaningful in its own right with regard to why a specific student pilot, or pilots, did or did not satisfy their instructor that they could fly solo, should be highlighted on the transcript”

Four issues led to more detailed extraction guidance. These were;

1. The quantity of information deemed to be a single critical incident. It was decided that a critical incident would be any singular phenomenon (behaviour) belonging to the student (action, inaction, skill, procedure, decision, communication etc) that the instructor was dissatisfied or satisfied with in terms of potential for solo flight, or any single reason given for not sending a pilot solo. In other words any individual unit of behaviour that was effective or ineffective with regards to attaining the general aim (Flanagan 1954).
2. The level of specificity or generality, in terms of defining what could be counted as a critical incident. It was decided that ‘general’ reasons were acceptable as long as a specific session, flight or student were referred to. General comments about why students should or should not fly solo were not counted as incidents.

3. It was decided that 'positive' incidents that were relayed within 'negative' anecdotes could be included and vice-versa (e.g. an instructor mentioning a reason why they prevented a student going solo last week, within an anecdote related to why they did fly solo 'this' week)
4. On occasions two critical incidents were identified within a single interview (e.g. because the instructor was repeating it for emphasis). These were only counted once, and con-joined if more could be gained.

The full extraction guidelines are shown in Appendix J The final working definition of a critical incident was:

> Any observed behaviours of the student pilot that were effective or ineffective in attaining the goal. In other words any unit of behaviour (action, inaction, decision, communication etc) that was meaningful in its own right with regard to the decision to allow or disallow solo flight. In addition to this any single reason given by the instructor as to why a specific student was or was not sent solo should also be included.

Incident extraction - Process

Critical incidents were 'extracted' from the remaining transcripts using the guidance in Appendix J), and the definition above.

Analysis and Categorisation of extracted incidents

As outlined, two independent sets of analyses were performed on the critical incident data.

Firstly, the 'meaning' analysis was performed using qualitative techniques based on inductive reasoning and principles of grounded theory (Strauss & Corbin 1967). This was the same process described in chapter six and utilised 'open-coding' and the same paired-comparison technique as described by Partington (2002). Concepts were kept deliberately narrow rather than attempting to cover large areas of meaning within single categories, in order that there was no loss of fidelity, since generalisation would take place with the higher levels of categorisation (axial coding). This led to a large number of lower level categories. Those that described similar overall phenomena were grouped to form larger categories (using the axial coding process), resulting in two further levels (termed second level and top level categories).

Secondly, template analysis as described by King (1998) was performed using the flight phase template from previous research by Jarvis and Harris (2008). Each critical incident was coded by the flight phase in which they occurred. Template analysis allows the researcher to form new categories and change or delete existing categories as the analysis progresses. The data therefore changed the flight phase template and only once finalised was it used to categorise the data from scratch.

Matching of CIT categories to causal accident factors.

The 59 bottom level categories from the human factors (HF) causal template (chapter 6, Table 6.1) were matched to the top level categories that emerged from the CIT 'meaning' analysis (as if each was a critical incident). Associated accident narrative examples were provided with each one in order to add context, as well as both template structures.

A second rater independently conducted the same analysis and got the same result, bar two of the 59 categories (both classified as airmanship rather than attention, and hence within the same top-level category of the accident analysis: 'Attention'). Hence there was no need for a statistical reliability test to demonstrate this high level of agreement. This showed that the two sets of analyses were the same in some areas. Results are presented in the results section.

Appendix K shows the results of this matching process.

Reliability and validity processes.

Six out of the eight processes detailed in Table 9.1 were completed fully. Two checks ('General SME review', and the 'participation rates' check) were attempted but found to be less applicable than anticipated due to issues relating to the specific context of the research. These were used in a limited way only. This section outlines all eight processes. The 'participant cross-checking interview' and the 'descriptive validity check' are discussed together because they were practically implemented as one process.

Independent incident extraction check

This was used as an integral part of the incident extraction process to code valid incidents from the narrative. An independent rater coded every fourth transcript after the initial four had been done jointly (meaning that just over 31% were assessed by both raters).
There is no general rule as to how to go about measuring such a check. It is not a simple matter of categorisation into known groups leading to a quantitative reliability check because the test has a qualitative element in the determination of meaningful events. Therefore no simple solution exists in terms of measurement.

Where raters disagreed on whether a particular piece of description should be extracted, the inter-rater analysis was straightforward (as long as some reasonable flexibility was allowed around the exact number of words highlighted as part of that event). However the most common issue of disagreement between raters was what number of incidents that a description constituted (one, two or three). Additionally, the amount of narrative highlighted as being an incident usually differed slightly between raters. These areas warranted much discussion.

In total, fourteen interviews were coded by both raters, yielding 102 critical incidents, only five of which could definitely be established as not being selected by both raters (representing 96% agreement on which parts of the narratives warranted 'critical incident status'). A third rater helped to resolve differences in terms of scoping incidents (i.e. what constituted a single incident). From the extraction check process it was clear that the coding of critical incidents was a reasonably consistent and reliable process as long as the guidance (shown in Appendix J) was closely followed.

Participant cross-checking interview and descriptive validity check

'Participant cross-checking' and 'descriptive validity' were employed in order to ensure that the interpretation of incidents by the investigator was in line with the intended meaning of the participants who supplied them. Eight instructors were re-interviewed about their incidents several months after taking part originally. Each was presented with their own critical incidents one by one (along with the original interview transcripts where necessary in order to give context and aid recall). The categorisation system was presented and explained to the instructors and the purpose of the study reiterated.

For each critical incident, the instructor was shown the initial low-level category label and asked whether they felt it was adequate to explain the phenomena that they had been trying to express. They were asked to look across all the low level categories and determine whether they felt that that critical incident fitted any other categories better. Over the course of the eight interviews, 74 out of 82 incidents were agreed upon by the instructors as being adequately classified, in terms of low level categorisation. This demonstrated that the categorisation process was satisfactory in terms of the analysis of meaning and initial categorisation of incidents. Participants were also asked to add, delete or amend incidents as they saw fit (in line with the 'participant cross-check'). In practice only a few decided to make any changes, and some wished to add a little detail. After some of the interviews, category titles were modified to better reflect the incidents.

During each of these interviews discussion was encouraged about the whole categorisation process. After each of these sessions the investigator reviewed the overall categorisation and made some changes where appropriate in line with the feedback. Additionally during this process, participants were also asked to say which flight phase the incident took place in or whether it was a general comment that did not refer to any given part of the flight. This information was used to help validate the categorisation during the template analysis of flight phases.

One of these interview sessions was overseen by a second subject matter expert who was closely involved with the project, in order to discuss the process afterwards and ensure that it was objective.

Categorisation by an independent judge (Reliability check)

Butterfield et al (2005) describe the process of categorisation by an independent judge as a sample of incidents being placed into existing categories. This check has previously been referred to as inter-rater reliability testing. It was done once the initial categorisation was complete at all levels.

An independent rater categorised extracted incidents into top level categories. This tested the entire chain of categorisation (from extracted incidents to top-level categories). The sample of critical incidents was drawn at random from all extracted incidents, by selecting every sixth incident (meaning a sample of 110). Each incident was presented within surrounding interview context where needed. At the time of this check there were 28 top level categories in total (13 negative and 15 positive).

A gliding SME with professional flying experience and human factors / CRM background was employed for this test. He was made familiar with the categorisations (all levels) and the objective of the process. Because the second rater had access to the mid level category titles, incidents could first be placed into one of these if it helped to decide on the higher level category.

The percentage of agreement was calculated on the basis of this test. There was agreement upon 96 of the 110 incidents (percentage of agreement 87.3%); no negative incidents were put into any positive categories and vice-versa. This gave a Cohen's Kappa value of 0.86 which was deemed highly satisfactory. The test data are shown in Appendix L, along with a table of 28 top level categories at the time of the test.

During this process notes were made by the investigator about any issues that appeared to arise, and following the process the categories were scrutinised in discussion with the independent rater. Following these discussions, several small adjustments were made to some categories as outlined in Appendix M.

Tracking the redundancy of incidents during data collection

This check was used in order to establish the point during data collection when it could be confidently claimed that the subject area had been exhausted and hence data collection could cease. The number of new categories emerging (from open coding) from each batch of about 50 incidents was tracked, in line with previous research (e.g. Bradley 1992).The point was reached where no new categories were appearing at all (either positive or negative). The processes used are described in detail in Appendix N, along with the graphs of results.

Because the numbers of incidents was not known until after the extraction was complete, the exact numbers of incidents (i.e. batches of 50) could not be planned for accurately and so the concurrent check was only a practical approximation. Once all data had been collected and categorised the positive and negative incidents were checked in batches of exactly 25, in order to uncover the pattern of redundancy more accurately. Figure 9.1 shows the progress of redundancy of negative incidents against positive incidents. It is of interest to note that the negative incidents took many more batches to reach the point of saturation due to higher numbers of new categories continuing to emerge throughout the process. This was partly because each interview provided fewer positive than negative comments.

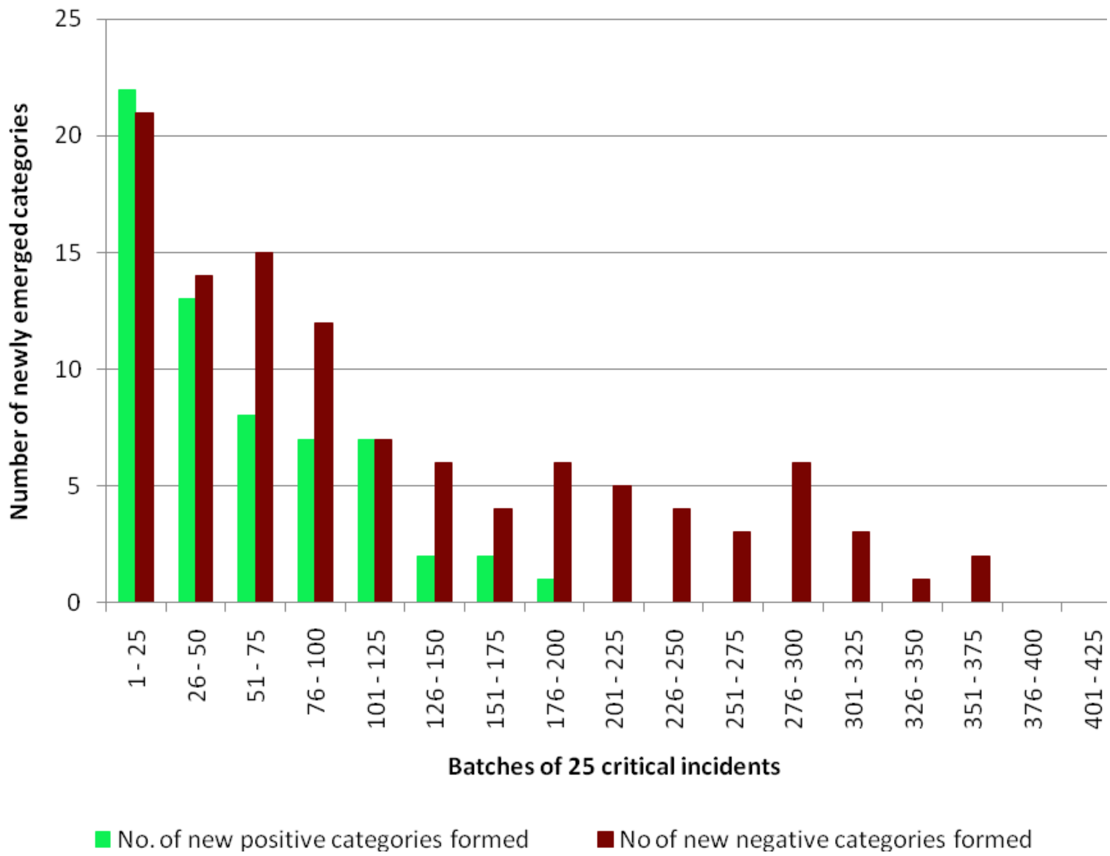


Figure 9.1. Number of new categories that emerged per batch of 25 critical incidents.

Figure 9.1 shows that after collecting about 225 positive comments and 400 negative ones, the point was reached where no new 'meaning' emerged from the collection of further critical incidents. Hence the subject area had been properly covered and data collection could be stopped.

General SME review

Because the participants were also the SMEs it was found that this process risked upsetting the objective nature of the data. For example, whereas most participants agreed with categories around the area of judgement and positioning, several advised that categories such as airspeed control and co-ordination should be removed on the basis that such basic skills ought to be attained long before a student was being assessed for solo. However this does not make them scientifically non-valid as data, nor does it mean anything about the validity of the categories. Indeed, since during the initial CIT interviews many instructors claimed that deficiencies in basic skills (poor control co-ordination and speed control) made them decide not to allow solo flight, it would be wrong to drop these categories on the basis of the opinions of other instructors. It became clear that asking instructors to comment upon the overall categorisation was not a valid activity in itself, and so this process was restricted to a few SMEs who were closely involved with the project and who had a good understanding of the aims, the process, and the data.

Participation rate check

Butterfield et al (2005) suggest checking the validity of incidents by establishing the proportion of participants who cited such incidents (at the first level of coding). A participant rate of 25% is suggested as valid. This assumes that unless a critical incident is put forward by a reasonable number of participants, it is not of value. However it is clear that this would depend on the specificity of coding. Incidents coded 'generally' would have less chance of possessing low participant rates. Since the present study used very narrow coding at the first level, this check risked losing important data. The check was therefore only used on the top-level categories. Despite this there remained a large number of positive categories that had very low participation rates (see Table 9.2). This could have been partly due to the lower number of positive incidents than negative incidents in each interview. Every category with a participation rate of less than 25% was scrutinised by the investigator and two subject matter experts. It was determined that the only category that could be justifiably dropped was 'unusual glider type'. This had the lowest of all participation rates, with only one critical incident. All other categories were deemed to valid by the SME group, albeit with low participation rates.

Interview fidelity

In order to ensure the proper rigour was being upheld in the interview process, an interview fidelity check was employed. A random sample of interviews (recordings) was listened to by an experienced researcher. The individual was a chartered psychologist with a PhD in applied psychology, extensive interviewing and qualitative research experience in aviation psychology, but was not linked to the project in any way. The researcher particularly checked that the interview technique was being upheld to the schedule and that participants were not being led in any way. No problems were highlighted in any interview.

Results

Firstly, the results from the 'meaning analysis' of critical incidents will be reported, secondly those from the flight phase analysis of critical incidents, and thirdly there is a short numerical comparison of these two sets of analyses.

Data

Forty-four instructors were interviewed. The mean total experience was 2,431 hours (SD = 2333 hours) although most instructors gave approximate estimates of experience. All participants held a British Gliding Association assistant or full rating, two were regional examiners and three were current professional instructors. Overall, the sample represented eight UK gliding clubs, including three ridge sites, two coastal sites, one mountain site, three clubs based on former military airports, two very large UK clubs and two very small clubs. Within this there were a variety of glider types, operations, airspace restrictions and traffic densities. Interview duration ranged from 14 minutes to 95 minutes.

Final Taxonomy of incident themes

There were 659 critical incidents after coding was complete. The final three-level taxonomy consisted of 23 top level categories (12 positive and 11 negative), 52 mid-level categories (24 positive and 28 negative) and 169 low level categories (61 positive and 108 negative).

Top level categories are listed with incident counts and percentages in Table 9.2. The mid level structure feeding into these is shown in Figure 9.2 (for negative incidents), and Figure 9.3 for positive ones. A full list of bottom level categories, numbered for identification across all levels is shown in Appendices O and P.

Table 9.2. Top level categories. The 'Ref' column contains the category reference. Column 'n' gives the total number of critical incidents under each category, and column '%-n' gives this figure as a percentage of either all negative incidents (432) or all positive incidents (277). Column 'pr' gives the number of participants who gave at least one critical incident in that category and '%-pr' gives this figure as a percentage of all 44 interviewees. These latter figures represent the participation rate of each category.

	Ref	Name of top-level category	n	%-n	pr	%-pr
NEGATIVE (Did **not** go solo because)	N 1	In-flight attention	**28**	**6.5**	**18**	**41**
	N 2	Behind the glider	**10**	**2.3**	**7**	**16**
	N 3	Handling and control	**85**	**19.7**	**35**	**80**
	N 4	Situation – Instructor disharmony	**10**	**2.3**	**9**	**20**
	N 5	Decision / Strategy	**30**	**6.9**	**18**	**41**
	N 6	Airspeed control	**51**	**11.8**	**25**	**57**
	N 7	Indirect issues	**44**	**10.2**	**24**	**55**
	N 8	Consistency	**13**	**3**	**9**	**20**
	N 9	Airmanship	**15**	**3.5**	**14**	**32**
	N10	Not Coping	**33**	**7.6**	**19**	**43**
	N11	Judgement and Positioning	**113**	**26.2**	**36**	**82**
POSITIVE (Did go solo because)	P 1	In-flight Attention	**6**	**2.6**	**3**	**7**
	P 2	Instant Reactions	**3**	**1.3**	**3**	**7**
	P 3	Handling and Control	**25**	**11**	**17**	**39**
	P 4	Situation-Instructor' harmony	**31**	**13.7**	**15**	**34**
	P 5	Decisions	**3**	**1.3**	**3**	**7**
	P 6	Airspeed control	**8**	**3.5**	**6**	**14**
	P 7	Indirect issues	**29**	**12.8**	**18**	**41**
	P 8	Consistency	**9**	**4**	**6**	**14**
	P 9	Airmanship	**4**	**1.8**	**4**	**9**
	P10	Dealing Well With a Simulated Situation / Emergency	**27**	**11.9**	**12**	**27**
	P11	Judgement and Positioning	**29**	**12.8**	**15**	**34**
	P12	Overall non-specific	**53**	**23.3**	**21**	**48**

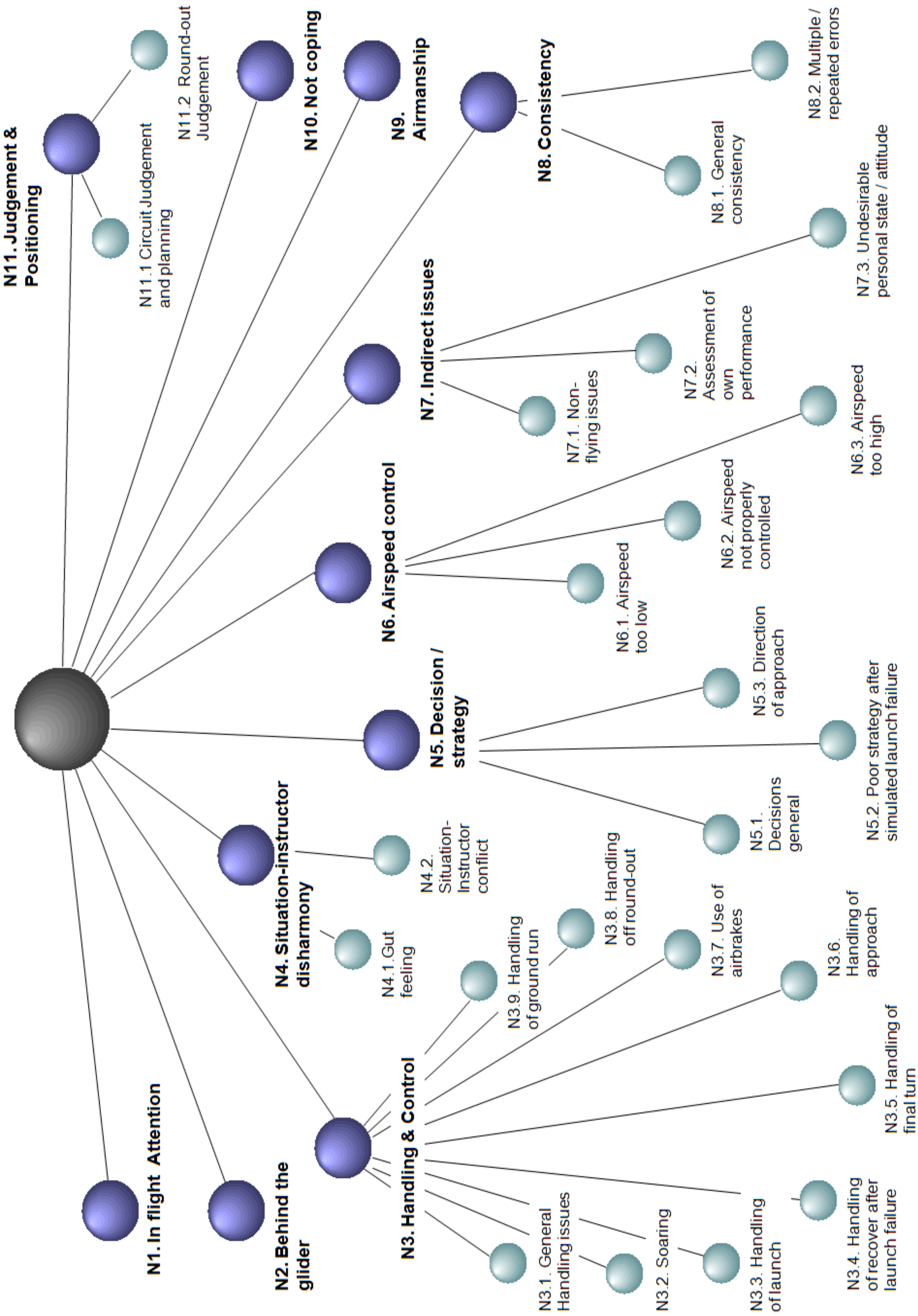


Figure 9.2. Top and mid-level categorisation of negative critical incidents

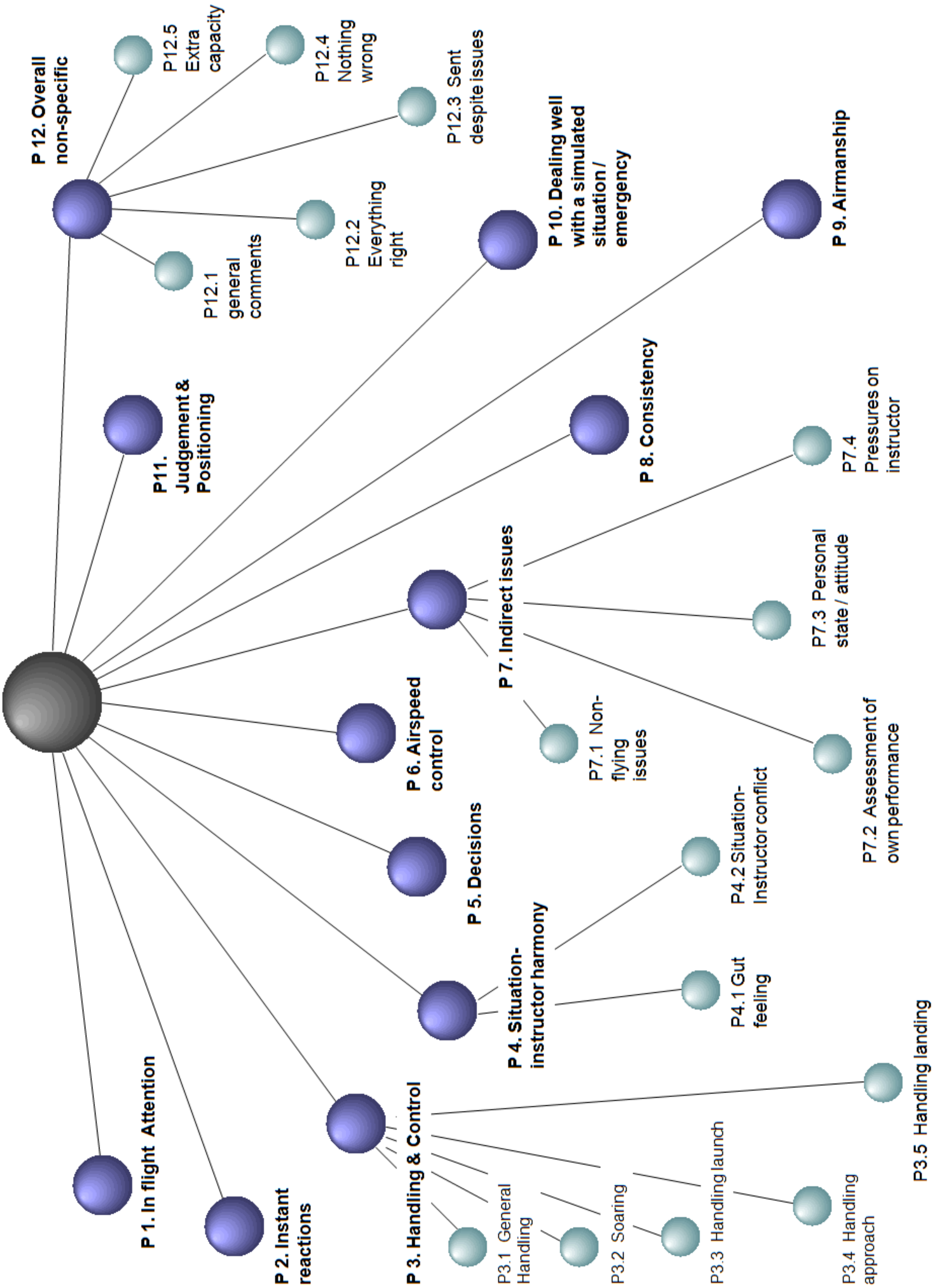


Figure 9.3. Top and mid-level categorisation of positive critical incidents

Top level categories

The most fundamental finding was the extent to which polarised pairs of top-level categories emerged. All top level negative categories had matching parallel positive top-level categories. Any standardisation of labelling between the category pairs was undertaken with advice from subject matter experts. The top-level category 'overall non-specific' was not able to be paired in this way.

Category Explanations

N1 and P1: 'In flight attention': These included all incidents where some form of attention was apparent as the primary factor during flight (as opposed to on the ground). This could be in the form of a control slip or attending to the wrong task (as perceived by the instructor), distributing attention improperly between tasks, or not being aware of all the aspects of the situation. For this reason, comments using the term "situational awareness" were grouped here. 'Situation awareness' can be simply defined as "knowing what is going on around you" (Endsley & Garland 2000) and hence expresses the same general concept as the other comments grouped into this category (positive or negative).

Examples:

- "oblivious to the fact that we were in crashing sink all the way down the downwind leg" (N1.1.7)

- "I asked him several questions like 'where's the airfield now?' and without turning his head he said 'it's directly behind me', 'it's under the wing', so he was spatially aware without actually physically saying 'well it's over there somewhere' (P1.1.1)

N2 ‘behind the aircraft’: This was a commonly expressed idea that related to instructors’ perceptions that the student was not reacting promptly, or waited too long before taking action. Terms such as “behind the aircraft” were common.

Examples:

- “They’re sort of reactive. They reacted to things, rather than actually be in charge... reacted when things started to go wrong. You would say he was behind the aeroplane.... As an instructor, you know that because you want the stick to move well before it does, and you feel that everything happens after you want it to.” (N2.1.1

- “They did not react at the rate that I would have wanted them to react” (N2.1.2)

P2: Instant reactions: This was related to specific, discrete incidents where the student was said to react quickly and properly (commonly to sinking air).

Examples:

- “Instant reaction to the sink that you get over the Tring road... flying along and there’s this sshhhoo the nose goes down, turns in.. Just the reaction and the erm, the response to the unusual, that’s very good.” (P2.1.1)

- “I’d asked the tug pilot to wave me off at 1800ft and he didn’t hesitate, you know, tug come up to 1800ft, waggled it’s wings and he [gesture of pulling the release], ‘...what was all that all about?’, so he reacted accordingly” (P2.1.3)

N3 and P3: Handling and control: These categories contained all incidents that described aspects of glider control and handling, except those explicitly related to airspeed. These include the handling of specific tasks (e.g. launch rotation, ground run), general handling of the glider (e.g. control co-ordination), handling of specific flight phases (e.g. the final turn), and the use of controls such as airbrakes and wheel brakes. N3 was the largest of all categories, containing numerous sub categories and low level categories.

Examples:

- "pulled back on the stick to the back stop... not in a sharp stabbing motion... in a deceptively smooth consistent one.. rapidly accelerated into a very high G situation which I couldn't, I could barely recover from...An over steep launch" (N3.3.6)

- "Not allowing the aircraft to fly itself... moving the controls, and all the aircraft is doing is creating drag. Now that is poor technique" (N 3.1.8)

- "not snatching at the controls, not kind of jerking, don't know how else you can define 'smooth'", (P3.1.1)

N4: Situation – Instructor disharmony:

This included critical incidents where the participant described feelings of negativity (e.g. discomfort, unhappiness, uneasiness) towards the situation or the pilot's actions, or having to intervene as a result (prompting, taking control etc). Specific descriptions of what the student did (such as misjudging or mishandling) were not included here, since these were prioritised into specific categories.

Examples:

- "your gut feeling is saying you're uncomfortable about them, so you may modify your overall evaluation on that"(N4.1.1)

- "You know that feeling where you just want to take over all the way round, but you don't. It was like that, it never really felt right" (N 4.2.3)

P4 Situation - Instructor Harmony

The opposite to N4, this includes incidents where instructors had expressed the feeling that the situation (including the flying of the student) was close to (or at) what they felt comfortable with. This category was much larger than its negative pairing (N4), in both real and relative terms (n = 30, representing 14% of positive incidents).

Examples:

- "He did everything when I would have done it.. it's when the stick moves as if you were flying, and you know he flies like you would do. That's what this guy was like" (P4.2.1)

- “the experience of sitting in the back and feeling sort of useless, redundant, is the best way I can put it... You sort of get the feeling that you are superfluous to requirement” (P4.2.2)

N5 Decisions / Strategy:

This category was used where it could be determined that the decision-making or strategy in use was being referred to. Poor decisions and decision uncertainty were included. A very large group of these incidents formed mid-level category called ‘Poor strategy after simulated launch failure’. This is recognised as a critical decision making situation, because in a matter of seconds the pilot must decide between landing ahead and turning the glider, both of which demand other quick choices to be made subsequently.

Examples:

- “then the decision making, and you know, dealing with that was just not good... I think also the decisions being made that weren’t necessarily the best or the right ones... the decision making wasn’t as good as it should have been” (N5.1.1)

- “[circuit], just not really making his mind up… general sort of uncertainty about everything” (N5.1.2)

P5 Decisions (positive):

This category was very small, containing only three comments, each where good decision making was mentioned in a general sense, rather than as a specific instance.

Example:

- "He did the cable breaks, and yeah they were ok, yeah good, decisions... That's actually quite challenging here as you know" (P5.1.1)

N6 and P6 Airspeed Control:

Comments relating to airspeed control were coded here. These incidents were kept separate (i.e. from those of 'handling and control') because a large number of participants referred to airspeed control quite specifically and separately. Although inaccurate airspeed can be the result of poor handling, it can also be caused by other factors such as prioritisation of pilots' attention on a tracking task (Padfield et al 1998). Maintaining a separate category therefore avoided inappropriate inferences on the part of the raters. It is noteworthy that category N6 had a high participation rate.

Examples:

- "I mentioned low speed on the winch launch (SJ: What speed?) 50ish... well certainly going up the, the steep bit of it. I think the entry wasn't too bad and then as we went up into the steep part of the climb the speed was dropping off..." (N6.1.6)

- "until the final turn... suddenly the speed is up round the clock and we're coming in like a bat out of hell, sort of 70 or 80 knots" (N6.3.1)

N7 and P7 Indirect issues:

Critical incidents relating to issues not directly relating to flying the glider were coded here. These included issues such as the quality of pre-flight checks, trainee self assessment and personal state, and pressures encountered to send a student solo (such as the fact that he or she was 'overdue' to go).

Examples:

- "Worst thing, he was indignant about the whole thing. He said he'd done what I told him to [after the first flight] and it hadn't worked... bad attitude" (N7.3.1)

- "He was incredibly tense both physically and mentally, With Peter you push the stick and it was held in his iron grip. You'd also see the shoulders, you know, so there's physical tension which was a kind of reflection of his mental tension" (N7.3.4)

N8 and P8 Consistency:

These comments referred to repeated performance (or the inability to repeat performance). Although forming only small categories (in both the positive and negative sense), the concept of 'consistency' was specifically emphasised as being very important by a number of participants when making their decisions.

Examples:

- "from time to time he'd fly several perfect circuits, and then he'd get one completely wrong...so there wasn't the consistency" (N8.1.1)

- "She was consistently good on the day.. it's all about consistency" (P8.1.1)

N9 and P9 Airmanship:

A number of instructors specifically used the term 'airmanship', usually in the context of lookout or actions concerning other gliders. Although many authors use the term airmanship, few attempt to explain what it is (see Thom 1997, Campbell & Bagshaw 2002, Green et al 1991). Pooley (2003) explains airmanship as encompassing "all aspects of safety such as awareness, rules and regulations and matters which differentiate the 'good pilot' from the 'bad pilot'". This is similar to a definition by Pratt (1994) which deems it to be "the quality that differentiates a pilot from an aeroplane driver". These definitions appear to suggest that airmanship is not concerned with technical flying skills, but other necessary elements of knowledge and awareness. Several popular UK general aviation text books (Thom 1997, Pooley 2003) are composed of sections for flying exercises, most of which include sub-sections entitled 'Airmanship'. A common feature in all such sub-sections is the emphasis on 'lookout', and "good visual awareness of other aircraft". Hence airmanship appears to be a popularly used term in aviation, often used in relation to keeping good lookout and avoiding others. This is how it was interpreted for the present study.

Comments relating to lookout or specifically using the term 'airmanship' were categorised as 'airmanship'.

Examples:

- "his lookout... lookout on a ridge is particularly important, airmanship I should say airmanship rather than lookout, his airmanship was particularly good" (P9.1.2)

- "because their lookout was so aw', non-existent actually! I started counting and got to over thirty before he moved his head" (N9.1.3)

N10 Not coping:

Comments that a student 'did not cope' or 'could not deal with' the situation were often related to simulated emergencies or unusual situations. This category did not include those comments regarding the technical symptoms (e.g. poor handling co ordination or inaccurate speed control), but only comments regarding the instructors perception or whether the student failed to managed the situation. Specific descriptions of symptoms were categorised here such as "he froze on the controls", as well as general comments such as "..about coping with pressure, he hadn't got it". Instructors' perception about students' ability to manage emergency situations in the future was also included. Another set of comments related to students' inability to manage workload or multiple tasks.

Examples:

- "He just didn't demonstrate his ability to cope with more than one or two things happening at once" (N10.1.1)

- "he's not going to cope with something going wrong, or something changing, you know a situation, like other traffic, or a patch of sink in the wrong place" (N10.1.3)

P10 Dealing Well With a Simulated Situation / Emergency:

Many critical incidents cited examples of students dealing well with unusual or ‘emergency’ situations, such as simulated or real launch failures, or the instructor engineering a challenging circuit. This category was not concerned with technicalities such as accurate speed control (as these were categorised elsewhere) but more with general comments about the way the students dealt with the pressure or the limited time available. Hence this category was similar (in an opposite sense) to N10, although it only related to emergency situations, simply because that was what the data provided.

Examples:

- “I threw the book at this bloke... he coped, difficult cable breaks, actually a very gradual power failure... he just ripped through it all” (P 10.1.1)

- “simulated cable break, and they’ve done that, it’s not what they’ve been used to but it hasn’t phased them, they’ve sorted it.” (P 10.1.7)

N11 and P11 Judgement and Positioning

It was sometimes not possible to reliably separate comments about judgement in the circuit from those concerning circuit positioning (often called ‘circuit planning’ by glider pilots), since positioning may be based on more than simply judgement (e.g. extrapolation of height loss, knowledge of the wind conditions etc). The method of flying circuits is a critical skill in gliding, and not simply about judgement. Because of this, the category was called ‘Judgement and Positioning’ as opposed to simply ‘judgement’, in order to reflect comments about glider positioning where it was not possible to isolate judgement elements from other complex cognitive skills. Notwithstanding this, the categories N11 and P11 contained all judgement related comments, not just those applicable to circuit flying. Because there were no positive comments about round-out judgement (landing), only N11 consisted of two intermediate categories to divide circuit and landing comments. All comments relating to glider positioning were found to relate to the circuit.

Note: N11.2 contained comments about misjudgement of the round-out. These were concerned with rounding out too early, too late, too far or close to the ground. Handling issues in the round-out and issues of technique (such as ballooning after closing the brakes or snatching at the control column) were coded into category N3 (Handling and control).

Examples:

- “a problem with circuit judgement I think really... just couldn’t seem to judge the circuit properly...a mixture of too tight, too far out” (N11.1.7)

- “able to plan a good circuit and keep it just right” (P11.1.2)

- Didn’t have the judgement to hold off.. flying it into the ground .. generally the actual visual judgement wasn’t good enough

P12 Overall non-specific

A large numbers of critical incidents (on the positive side only) concerned students doing ‘everything’ right, or doing ‘nothing wrong’. This was different to ‘consistency’, which was about students doing one thing right every time (or not). Similarly, comments were also made that summed up whole sessions (or flights), which did not occur in the same way on the negative side. Therefore category P12 was formed in order to cater for comments that were made about the overall flights or sessions in a general sense, without referring to any specific phase, element, issue etc. All such comments were positive.

Examples:

- “the whole performance is usually, they’re on top of” (P 12.1.4)

- “He did everything right. That’s all I can say really. There was nothing wrong with his flying” (P 12.2.1)

- “I couldn’t find a reason not to send him solo, on the day... “

- “It’s no mistakes really. It’s more like what I don’t see, and if I don’t see anything wrong then that’s what indicates to me that they are ready. [SJ: And in this case, the guy you were thinking of?]. Yeah yeah, that’s like it, I didn’t see any mistakes” (P12.4.2)

- “they’ve got a little bit of spare capacity, to look around, you know. Asked him ‘where’s the airfield ?’, ‘it’s behind us on the left’, you know so enough spare thoughts over and above and that’s what I’m looking at, and this guy, in particular, was actually very very good... [they were in control of the aeroplane]” (P12.5.1)

Comparison of relative critical incident frequencies by category

(Figure 9.4) shows the percentage of negative critical incidents in each top level category, and the percentage of positive critical incidents in each. Using a percentage breakdown over categories enables direct comparison of positive and negative pairs in terms of their representation relative to other positive or negative categories. The percentage and total figures are shown in Table 9.2.

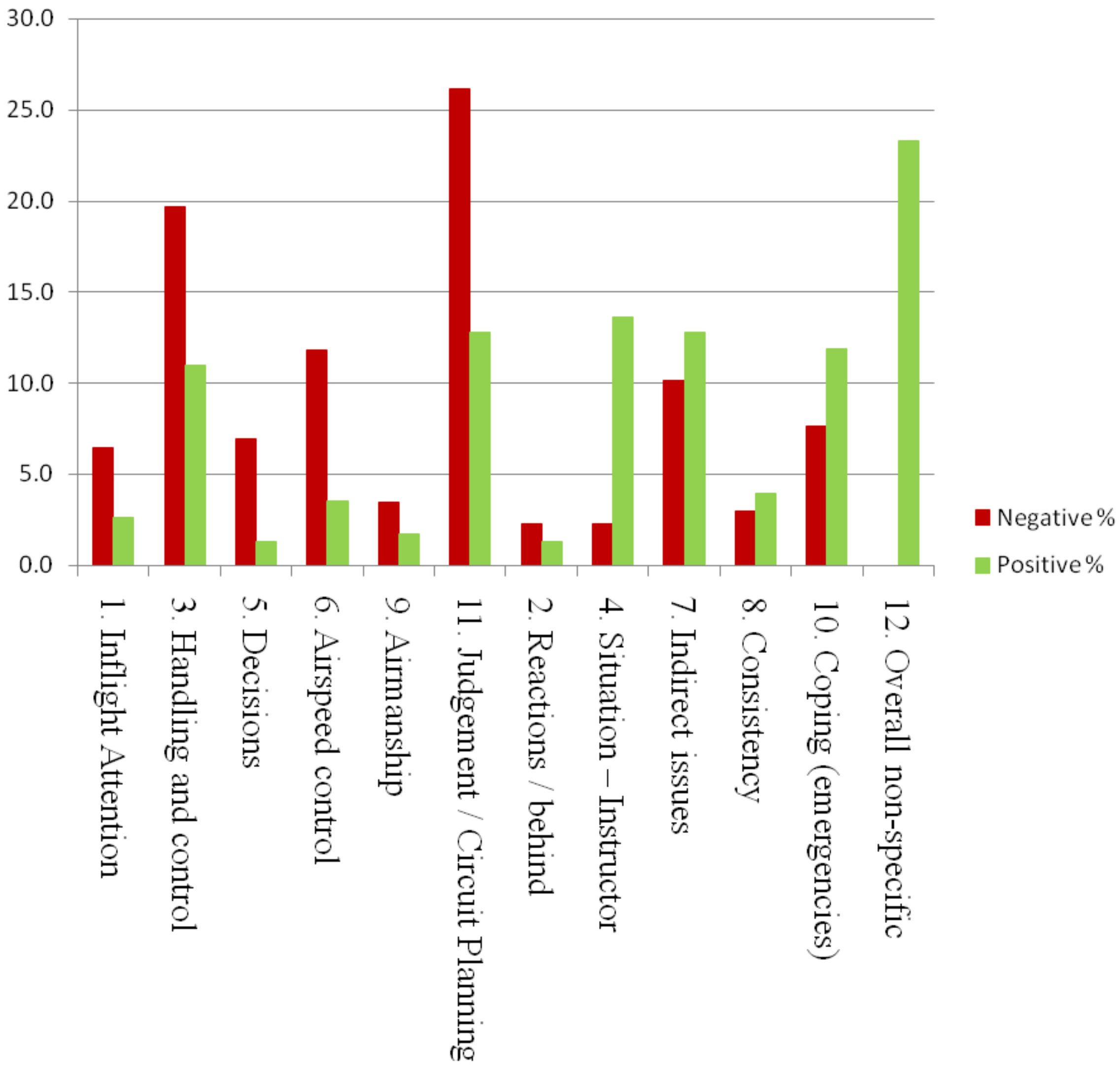


Figure 9.4. Percentage of positive critical incidents in each top-level category alongside the percentage of negative critical incidents in each.

When the 59 human-factor causal accident categories from chapter six (Table 6.1) were matched to the 23 CIT categories, six negative categories accounted for all of them. These categories were N1, N3, N5, N6, N9 and N11. These (along with their positive pairing) are placed on the left-hand side of Figure 9.4). The reason that only negative categories were used was because the causal categories were descriptions of accident causes. The issues themselves were equally relevant to the positive pairings of these six negative categories (since they describe the same issues, just with a different instructor decision outcome). These six negative categories and their positive counterparts were termed 'causal-factor' critical incident categories, because they described the same phenomena as causal accident events.

Table 9.3 shows how the four top-level categories established from accident analysis (chapter 6) map onto the six applicable pairs from the CIT analysis, along with the equivalent percentages of accidents and critical incidents. Airspeed issues were coded under the overall heading 'handling' in the accident analysis, whereas this was not the case for the CIT data. Hence, N6 (airspeed control) as well as N3 (handling and control) were required to account for all the low level accident codes within the overall 'handling' category from the accident analysis. The same was true of the attention category from accident data and the CIT categories of N1 (In-flight attention) and N9 (Airmanship). This is shown in Table 9.3.

Table 9.3. The ‘causal factor’ critical incident categories against the four top-level accident causal factor categories from chapter six.

Four main categories from HF accident analysis (Chapter 6)	% accidents (All pilots)	% accidents Inexperienced Pilots (10 hours and under)	Equivalent categories from CIT analysis	n	%
J. Perceptual judgement	24	53.6	**N/P11 Judgement, Positioning**	142	35.8
H. Handling	25	32.1	**N/P3 and N/P6 'Handling and Control' and 'Airspeed control'**	169	42.6
S. Strategy	26	0.0	**N/P5 Decisions**	33	8.3
A. Attention	25	14.3	**N/P1 and N/P9. 'In-Flight Attention' and 'Airmanship'**	53	13.4

The six ‘causal factor’ pairs (12 categories) accounted for 397 critical incidents (60% of the total), see Table 9.4. This implies that 60% of critical incidents reported by instructors as being influential in the ‘solo-decision’ process were descriptions of the same events found to cause accidents to UK glider pilots.

Figure 9.4 shows that the six causal-factor category pairs are all characterised by high percentages of negative incidents, rather than positive ones. The remaining categories (on the right of the graph) are all characterised by higher percentages of positive comments (reasons why pilots were sent solo), with only one minor exception. In total 322 (81%) of these causal factor critical incidents were negative (led to solo flight being disallowed) and only 75 (19%) positive (led to solo flights). In terms of raw incident numbers, this dichotomy is statistically significant to a very high level of confidence (see Table 9.4).

Chi-square analysis showed that there was a highly significant difference in terms of critical incident numbers in 'causal-factor' categories between positive and negative groups (χ^2 = 107, df=1, $p < 0.001$, see Table 9.4).This showed that a significant difference existed between positive and negative incidents, and that negative incidents were significantly more likely to be associated with the causal factor categories.

Table 9.4. Total number of positive critical incidents belonging to the six 'causal-factor' categories, against the number of negative critical incidents in those categories. Data used for chi-square analysis.

	Total number of incidents in all six **'Causal-factor categories'** (P/N 1, 3, 5, 6, 9 & 11)	Total number of incidents in all **other categories** (N/P 2, 4, 7, 8, 10 & 12)
Positive incidents	**75**	**152**
Negative incidents	**322**	**110**
Total	**397**	**262**

Flight Phase analysis

Each of the 659 critical incidents was coded by the flight phase they related to, requiring some adjustments to the original template (Jarvis and Harris 2008) including the removal of several fine-grained phases which were unused. No such phases were associated with accident causes for low-hours pilots from 2002 - 2006 inclusive (Table 5.9, Chapter 5). Table 9.5 shows the totals.

Table 9.5. Adjusted flight phase template and total number of critical incidents in each phase

	Phase		All Incidents		Positive Incidents only		Negative Incidents only	
	1	pre-flight		**10**		**2**		**8**
2. Launch	2	Launch (non-specific)	9		4		5	
	2c	Launch (Rotation)	6		0		6	
	2d	Launch (established climb)	9	**46**	0	**7**	9	**39**
	2e	Launch (Aero-tow cruise)	6		0		6	
	2f	Launch (release)	3		1		2	
	2g	Launch (Recovery to speed)	13		2		11	
3.General Flying	3a	General (general flying)	26		3		23	
	3d	General (ridge)	4	**37**	2	**12**	2	**25**
	3g	General (unusual manoeuvring)	7		7		0	
4. Circuit	4a	Circuit (join)	6		0		6	
	4b	Circuit	156	**208**	31	**32**	125	**176**
	4c	Circuit (abbreviated c/b)	28		1		27	
	4d	Circuit (final turn)	18		0		18	
5. Approach	5a	approach	13		2		11	
	5b	approach (post launch failure)	7	**20**	0	**2**	7	**18**
6. Landing	6a	Landing	54		12		42	
	6b	Ground Run	1	**55**	0	**12**	1	**43**
7. Non-Specific	7a	Non-Specific	267		149		118	
	7b	Un-specified launch failure	16	**283**	11	**160**	5	**123**
		TOTALS		**659**		**227**		**432**

Table 9.5 shows that 283 incidents were not sufficiently ‘specific’ to be coded by flight phase (section ‘7’ - ‘non-specific’). Unless the reasons given by instructors for their decisions were sufficiently specific to establish which flight phase they related to, they could not be coded into the six flight main phases. Prompts from the interviewer helped to ensure that the phase was elicited, but numerous critical incidents were still of this more ‘general’ nature. For example, many referred to overall aspects of pilot performance, or non-flying cues picked up from the student, for example;

> “I just felt like he was reacting to the aircraft…. Everything he did just seemed late” (coded as N2)

> “General mannerisms about the guy. When you can talk to a guy... and you know just by talking to him ‘this guy is switched on’ and he’s not going to be a problem... he’s going to do the business.” (coded as P7)

Reliability tests showed that such incidents could be consistently recognised (see method section and Appendix Q).

This had wider relevance to the overall analysis. All incidents could be broken into two types; specific (able to be recognised as relating to a specific element of the flight, and hence able to be coded by flight phase) and non-specific (too general to attribute to a particular flight element). ‘Specific’ incidents could be coded into flight phases because they usually described events, or symptoms of events. However, instructors’ observations of overall behaviour, performance, attitudes and even their own feelings had to deemed ‘non-specific’. Hence the flight phase analysis indicated a more far reaching dichotomy within the data set, which proved to be useful in explaining other results.

The only other subsequent issue relating to this was a subset of incidents termed ‘launch failures in general’. These occurred when instructors commented about ‘launch failures’ without referring to a specific phase within that failure, for example:

> "I did, awkward straight ahead [simulated cable break]... Coped perfectly, flying perfect. Also it had been a complete surprise, I'd talked about doing circuits with him etc"

Hence these were unable to be categorised by a flight phase and were deemed 'non-specific'. For readability, all the non-specific critical incidents are henceforth referred to as belonging to a 'non-specific' category; flight phase category 7, even though they do not really form part of the flight phase analysis.

Only 376 out of 659 critical incidents (57%) were of a sufficiently specific nature to allow categorising by flight phase. Notably however, whereas 309 out of 432 (72%) negative critical incidents were specific in nature, only 67 out of 227 positive comments were (just under 30%).

Table 9.6. Total number of positive and negative critical incidents coded as specific flight phases and non-specific flight phase. Data used for chi-square analysis.

	Total number of incidents coded into **Specific flight phase categories** (Flight Phase 1, 2, 3, 4, 5, 6)	Total number of incidents coded as **Non-Specific** (Flight Phase 7)
Positive incidents	**67**	**160**
Negative incidents	**309**	**123**
Total	**376**	**283**

The data in Table 9.6 was subject to a Pearson's chi-square test of association, resulting in a chi-square of 107.2 (df = 1, p<0.001). This showed that negative incidents were significantly more likely to be 'specific', and positive incidents were likely to be associated with general comment.

Because similar results had also been obtained for the six pairs of causal-factor critical incident categories, a Cohen's Kappa test of agreement was conducted between incidents that could and could not be coded into flight phases (specific and non-specific) and those that were and were not coded as causal factors (i.e. categories N/P 1, 3, 5, 6, 9, 11). The data are presented in Table 9.7. The result was a Kappa of 0.59 (a 'fair' level of agreement [Robson 2002], although very close to the 'good' bracket). This showed that the critical incidents that were able to be coded by flight phase also tended to be coded as 'causal-factors' (categories N/P 1, 3, 5, 6, 9, 11), and 'non-specific' critical incidents tended to be those that were not 'causal factors'. This meant a high level of association between critical incidents coded as 'specific' (flight phase) and those coded as a causal accident factors (from the CIT 'meaning' analysis).

Table 9.7. Total number of causal-factor and non-causal factor critical incidents against specific and non-specific flight phases. Data used for Cohen's Kappa test of agreement.

	Specific (flight phase coding)	Non-specific (flight phase)
ALL 'Causal Factor' Incidents (N/P 1,3,5,6,9,11)	321	76
ALL NON-causal factor incidents (N/P 2,4,7,8,10,P12)	55	207
	376	**283**

Comparisons of flight phase analysis with findings from accident data.

For comparison of figures with previous flight phase analysis, critical incidents in the non-specific category were dropped. This left 376 critical incidents which were specific to particular flight phases (67 positive and 309 negative).

Comparison was conducted using relative frequencies (percentages). The percentage of accident causal events occurring to pilots with 10 hours or fewer in each main flight phase was used for comparison, from Jarvis and Harris (2008). Figure 9.5 shows these compared against the percentage of critical incidents coded into each main flight phase.

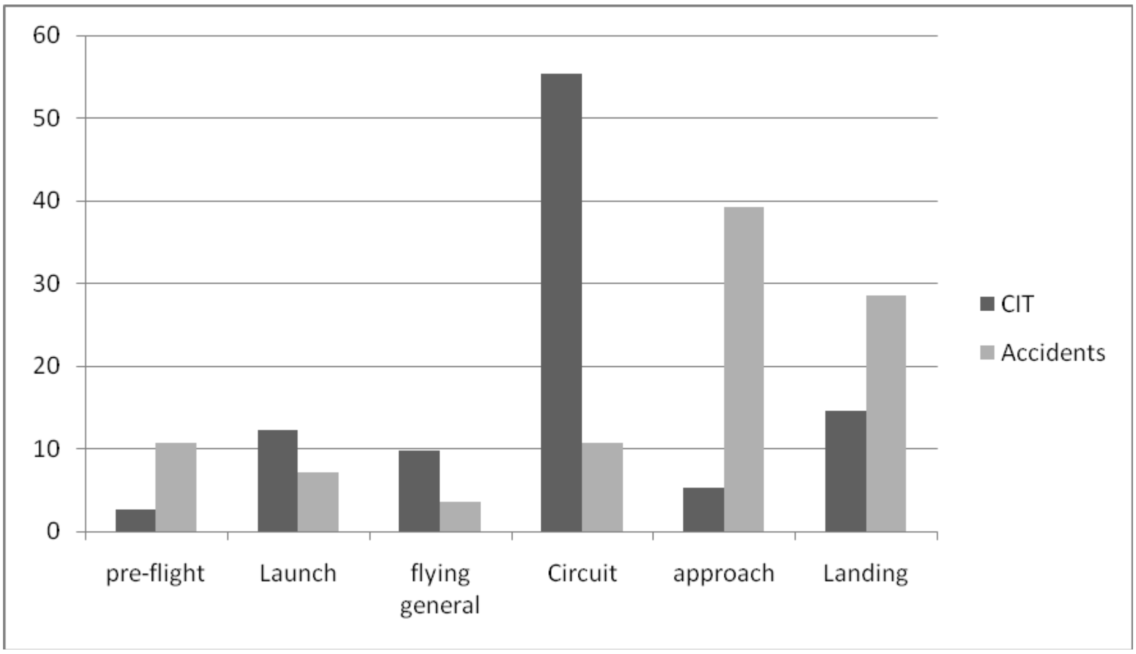


Figure 9.5. Percentage of critical incidents across the six main top-level flight phases compared to the percentage of accidents to pilots with 10 hours experience or fewer, split by the same six top level phases (as percentages).

Figure 9.5 shows that apart from 'pre-flight' the approach phase was the least commonly mentioned as influencing the solo decision process. The circuit was a feature of well over half of all 'specific' critical incidents, and over a third of all critical incidents. However the circuit was associated with the causes of very few accidents according to the data from Jarvis and Harris (2008).

In order to compare the present flight phase data with the data from chapter eight pertaining to instructor perception of accident likelihood, the flight phases were ranked by the total number of critical incidents that involved them. Figure 9.6 shows the ranked data against the instructor perceptions of the flight phases most likely to be associated with the causes of accidents for inexperienced pilots. 'Objective' rankings are also shown (from accident analysis, see chapter 5) using broken lines.

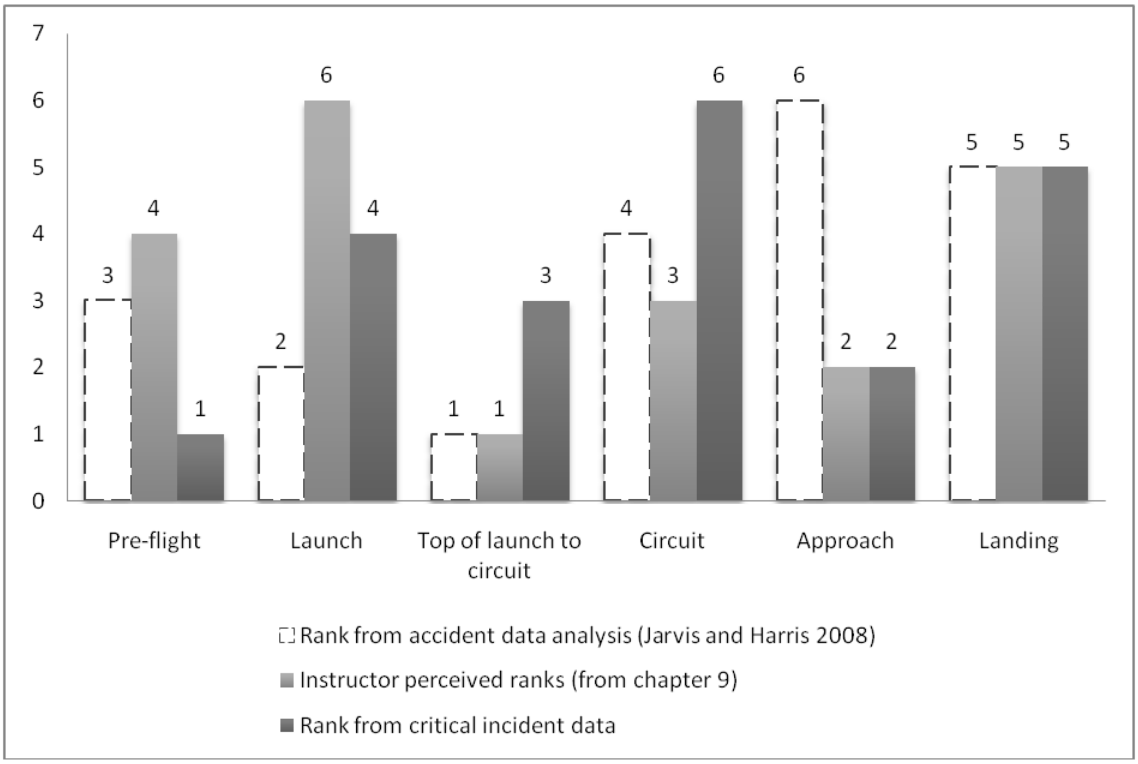


Figure 9.6. Ranked flight phase data (from critical incident numbers) against instructor perceptions of accident cause by flight phase for early-solo pilots (from chapter 8). Broken-lined bars represent the flight phase ranking from accident data ('objective ranking').

Numerical comparison of causal factors by flight phase results.

Table 9.8. Numbers of flight phase critical incidents split by causal factors

		Flight phases				
		Launch	General	Circuit	Approach	Landing
Causal Factor Categories	Judgement & Positioning			114		23
	Handling and Control (including Airspeed)	41	14	39	5	23
	Decisions and strategy			18	12	
	Attention (including "airmanship")	1	15	16		

In order to compare the results of the two analyses, flight phases were split by critical incidents fitting into one of the four causal factor categories from chapter 6. This was done to identify whether any particular causal elements dominate instructors' observations in specific flight phases, and enable further comparison with previous findings. Comments relating to the launch almost exclusively relate to handling and control. Circuit flying was dominated by judgement and positioning issues, and landing issues split evenly between judgement and handling. When referring to approach incidents, instructors predominantly mention issues of decision making and strategy.

The pre-flight phase was dropped from this analysis due to the very low number of critical incidents.

Discussion

This section has two parts, one relating to each of the general research questions:

1. What factors do gliding instructors identify as being satisfactory and unsatisfactory performance in students, in terms of allowing them to fly solo? This discussion draws mainly on the CIT 'meaning' analysis.

2. How do these findings compare to previous findings from accident data (in terms of areas of concern)? This is a discussion draws mainly of the flight phase analysis.

Some factors apply more generally and are therefore discussed across both sections. Section one explores the decisions made by instructors regarding solo flight, and attempts to model the decision process within the limitations of the data. This relies heavily on the findings from the 'emergent' CIT analysis. The second section explores how the findings relate to the sorts of accidents that glider pilots have been shown to sustain after they have recently been sent solo. This relies predominantly on the analysis of 'specific' incidents; i.e. those coded by a flight phase.

Section 1 - Instructor decision 'modelling'

This section will assess the findings of the critical incident 'meaning' analysis, and build an initial model of the decision attributes (Figure 9.8). The negative outcome (disallowing solo) will be treated first. The positive outcome will be explored subsequently (allowing solo) and the model presented lastly.

The model was based on evidence from the analysis and supported by ongoing SME discussion throughout the process to help with validation. It does not attempt to model the dynamics of the decision, only those attributes where some evidence was available. It is stressed that the model is hypothetical, and brings together the findings from the data as opposed to being put forward as a decision algorithm for which more research would be required. Further study would also be required to validate the model in a naturalistic setting. The manner in which many of the components of the model were put together is described concurrently throughout the discussion. The 'top half' of the model is presented first (negative outcome - disallowing solo flight) and a full diagram of the model, overview and explanation is at the end of this section.

Three key overall findings are fundamental to the discussion. Firstly, the finding that 22 of the 23 top-level categories could be matched into pairs of one positive and one negative category. Secondly, that when placed against the causal features of gliding accidents (from chapter 6) six of those pairs could be easily accounted for. Lastly, that the incidents comprising those six pairs accounted for 75% of negative incidents but only 30% of positive ones. This gives the rationale for splitting this section into two parts; 'disallow solo flight' and 'allow solo flight'.

The pairing of categories is less surprising when one considers that the nature of the decision is a dichotomy, either the pilot goes solo, or not. Against this background it would perhaps be more remarkable if such dichotomies did not emerge. Each of the 11 pairs is clearly of this nature, i.e. one member of the pair is the compliment of the other in terms of the decision process. About half the pairs are dominated by negative incidents and half by positive incidents.

Most causal accident factors from the findings of Chapter six were accounted for within those pairs that were dominated by negative outcomes. Again, this does not appear to be surprising. Most critical incidents within these six pairs were also independently coded as being 'specific' in nature, meaning sufficiently specific to code into one of six flight phases. These critical incidents took the form of clear and often discrete events occurring during 'assessment' flights. Positive incidents were predominantly unable to be coded in this way. Therefore, in contrast to the reasons given against solo flight (negative), the reasons given for sending a pilot solo were mainly of a non-specific nature. Overall this suggests that the decision outcome to prevent a pilot from going solo is not simply the compliment of the process used to allow solo flight.

Disallowing solo flight

As previously discussed, many aviation textbooks make reference to the standards needed in specific performance areas, e.g. “The usual standards apply to your take off, circuit and landing” (Thom 1997), and “the student would have perfected their landing technique” (Pooley 2003). This suggests that an important aspect in an instructor’s decision is whether the student is technically proficient in certain important areas of flight. In the present research the ‘important’ areas would include those with the most potential to cause an accident (i.e. the judgement and handling of the approach and landing). This section will describe the first part of the decision model (Figure 9.7).

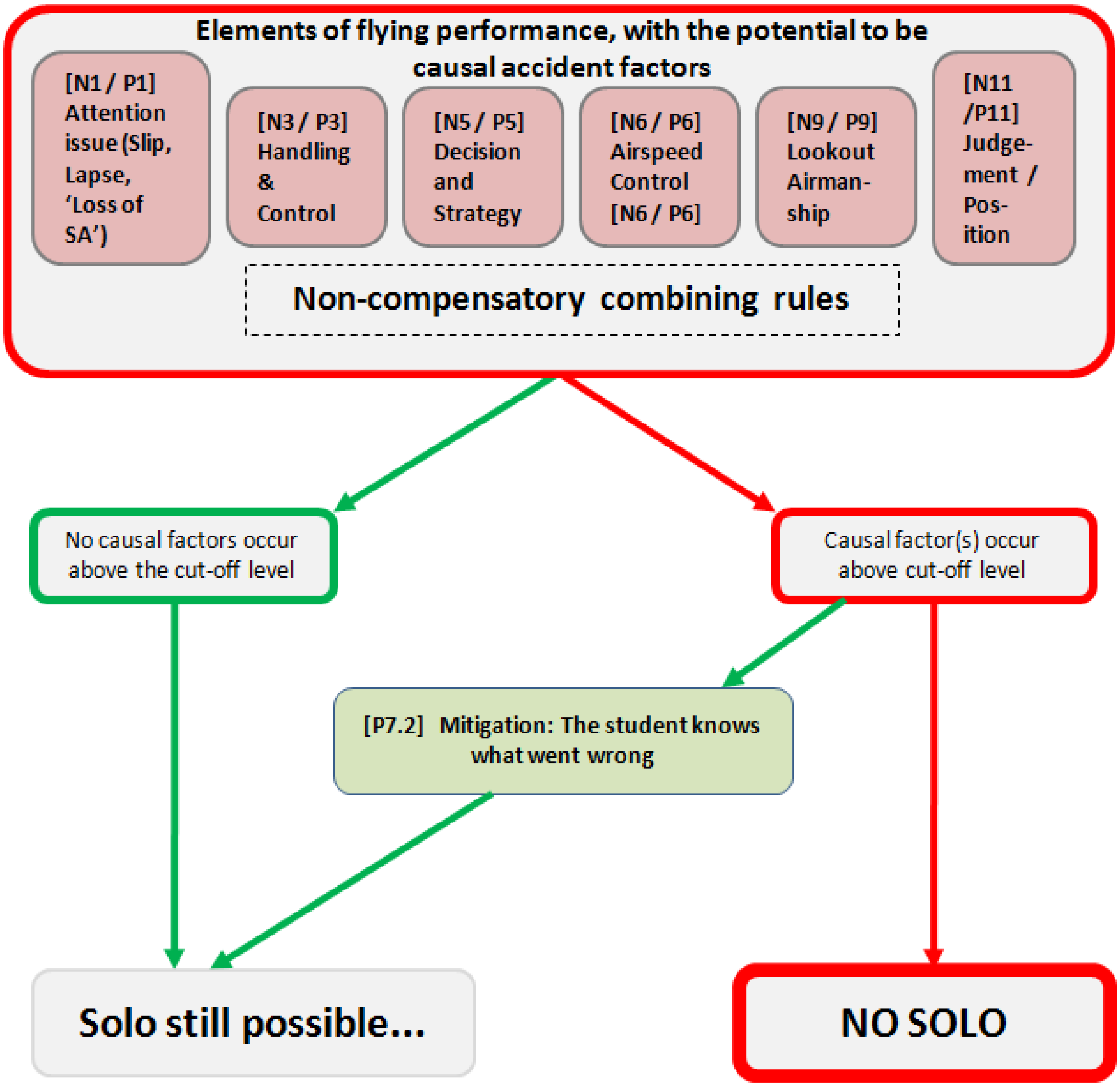


Figure 9.7. The decision model (intermediate stage) representing the non-compensatory decision process. As well as representing causal accident factors, the six areas (N/P 1, 3, 5, 6, 9, & 11) are all specific and observable as being within certain flight phases (as demonstrated). If no causal factors are observed (left side of model, in green) then the student may still go solo (the decision is still open). However sufficiently poor performance in one of these areas (sufficient to report as a critical incident) normally leads to the decision not to allow solo flight, in a non-compensatory way (right-hand side of the model, red lines). The only 'mitigation' that can compensate for such an event is demonstrated by category P7.2, which can keep the decision 'open' despite a causal incident having occurred.

The statistically significant result that most 'causal factor' critical incidents were negative (no solo) and most other incidents were positive (see Table 9.4) has possible implications for the process by which instructors weigh these types of evidence. Whereas instructors commonly pointed to the presence of one or two specific and observable events (termed 'causal accident factors') as reasons for disallowing solo flight, few mentioned any specific causal events (or absence of them) as reasons to allow solo flight. This indicates the possible existence of a simple 'discounting' type of decision model whereby the student cannot prove s/he is ready for solo, only disprove it by doing things wrongly. This would be a non-compensatory process since good performance could not compensate for bad. Category P12 ('overall non-specific') contains two subcategories that indicate this process from the positive incidents. These are called "everything right" (P12.2) and "nothing wrong" (P12.4). These are comments in which instructors explain that they did not see any errors, unsafe events or performance problems, and this contributed to them sending the student solo. P12.2 and P12.4 are clearly important categories because they contain 33 critical incidents in total (having two of the largest of all 62 low level positive categories). They are also further supported by similar ideas expressed through incidents in category P4 (situation-instructor harmony) such as 'Instructor not having to do or say anything' (P4.2.4) and 'Instructor was not worried / uncomfortable / uneasy etc' (P4.2.5).

This suggests a process by which instructors' base their decisions on whether or not they have seen something which could potentially cause an accident. An example of a common type of comment to support this would be:

"*there is nothing that's actually dangerous, you know, so off you go*" (P12.4)

From the CIT data, it is clear that such 'causal factor incidents' occur during many instructional flights, any of which could prevent solo flight being allowed by the instructor. There is strong evidence (from category P12) that pilots are sent solo only if no such incidents occur (and hence not sent solo if any occur). This suggests that for many instructors the occurrence of one such 'causal incident' is sufficient to prevent solo flight, and cannot be compensated for by the student doing well in other areas. No positive comments put across the idea that problem events could be 'forgiven' in the light of high performance in other areas. There is therefore no evidence that poor performance in some areas can be compensated for by good performance in others, and there is strong evidence (from the 'causal factor' incidents and from category P12 supported by P4) that no such compensation takes place, at least with regards to specific incident occurrences (potential accident causes, see table 9.7).

Such a process indicates a 'non-compensatory' method of combining attributes in the 'solo decision' process, in line with Ikomi and Guion (2000). The decision attributes that feed into the eventual judgement are represented by the six negative 'causal factor' incident categories (N1, N3, N5, N6, N9, N11). The findings of Ikomi and Guion (2000) showed consistent evidence that instructors ruled out solo flight if a student attained a score of less than three out of five on any manoeuvre in the syllabus. This implies that the non-compensatory rules were combined with a cut-off level on each attribute which might indicate the possibility of satisficing (because these 'cut-offs' could indicate aspiration levels). This may be what Thom refers to as "the usual standards... in take-off, circuit and landing" (Thom 1997). This would suggest a conjunctive form of non-compensatory decision model. This is where the decision maker sets certain cut-off points on various decision attributes and no attribute falling below the cut-off is viable (Hogarth 1987). However because critical incidents were collected rather than numerical scores, such process detail cannot be verified by the present study. A conjunctive non-compensatory model would imply that students do not have to be perfect, simply good enough (satisfactory) in all areas. Some critical incidents were coded under the heading 'despite problems' (P 12.3.1), meaning that the instructor sent the person solo whilst recognising that some aspects were not perfect, e.g. *"there are some problems there but they're not major..."*. This would indicate that a 'cut-off' point could exist, as opposed to the decision being compensatory, since no attempt is made to justify poor performance in one area with good performance in another. There were no comments indicating that compensatory methods existed within the six causal factor areas. Since all 'causal factor' critical incidents indicated a contribution to the decision process, it is assumed that the mention of the incident by the participant meant that a satisfactory level had been achieved (for positive incidents) or not (for negative ones).

Although the analysis appears to show that causal accident factors cannot be compensated for, one possible compensatory element was found. It was in line with guidance that to fly a glider solo, a pilot should "be aware of his limitations" (Piggott 1997). Some critical incidents showed that when the instructor perceived that the student knew what they had done wrong, this compensated for the mistake itself (e.g. by correctly assessing poor performance in a de-briefing). Category P7.2 contained a suitably large number of examples demonstrating this effect, including:

> *"we had a discussion about erm final turn heights and approach... the minute I started talking about it ... they pretty much said to me what I was going to say to... they said: "yeah I know" and then they pretty much said.. So it was this thing about: if you know what you're doing wrong then you're safe. So they'd known almost immediately what had gone wrong"*

> *"...It wasn't something they'd done on their first flights, so it wasn't a trend...<u>ok, you know what you did wrong there, I don't think you'll" [do that again]</u>*

This demonstrates the advantage of using an exploratory technique, since such a decision attribute could not have been uncovered using pre-determined dimensions such as used by Ikomi and Guion (2000).

Figure 9.7 summarises the findings of the decision process so far. Enough evidence is available to establish that six specific (and observable) areas relating to known accident causes are used as decision attributes by instructors (particularly with regard to preventing solo flight). Further evidence from positive categories (P12 and P4) in combination with previous research findings (Ikomi & Guion 2000), strongly suggests that non-compensatory combining rules are used by instructors. It is also possible that cut-off levels (or aspirational levels) operate as part of this process. Additionally, there is evidence of a separate process by which problematic events can occasionally be compensated for. This is a process in which instructors allow student knowledge and acceptance of the problem event to compensate for the event itself, and hence allow solo flight. There is no suggestion that this happens regularly. More research would be required to establish the extent to which this occurs.

Allowing solo flight

The process so far described, accounts for the vast majority of negative critical incidents (75%) and all major 'negative' type categories.

The process does not account for the categories that were most representative of the positive critical incidents (i.e. why solo flight was allowed). It has already been established (from the flight phase template analysis as well as the causal factor breakdown) that positive critical incidents were heavily characterised by their non-specific nature (see Tables 9.6 and 9.4). These categories were in line with advice in textbooks regarding the importance of non-technical and personal attributes in student pilots, such as "an honest and mature outlook" and the "ability to cope with emergencies" (Piggott 1997). This suggests the strong possibility that the types of evidence that instructors base positive decisions on (to send a pilot solo) are different to the evidence that they base negative decisions on (disallowing solo flight).

It has been established that most 'specific' critical incidents were related to non-solo decision outcomes, but that instructors tend to refer to general reasons when describing why they did send a pilot solo. It is therefore possible that to send a pilot solo, instructors need certain further evidence in addition to a lack of the negative evidence that they base non-solo decision outcomes upon. In other words, whereas an unsafe flight is sufficient to disallow solo flying, a safe flight is not sufficient to allow it. This would imply that the decision not to send a pilot solo is often easier than deciding to send them. If by noticing a single observable event a decision can be made (i.e. do not allow solo flight), then the instructor need go no further with the decision process. In this way Figure 9.7 may represent enough of the decision process to account for most non-solo decision outcomes, but only the first stage of the process required to send a student solo.

The nature of the 'non-specific' categories supports this. Each will be discussed in the following section, along with information relating to the further construction of the decision model.

Overall non-specific (P12).

This was the largest positive category, and exclusive to positive incidents. Much of the content has been discussed as giving support to the non-compensatory decision process. In this way, such instructor comments are not independent decision attributes in themselves, but comments about the decision process (and combining rules). For example, claims that "the whole performance... they're on top of" (P12.1), "he did everything right" (P12.2), "He didn't do anything wrong" (P12.3) are all comments about other decision factors, and not therefore independent decision attributes. It seems that many of the 'incidents' in P12 are in fact attempts to express the results of the decision process. Hence, with a notable exception (sub category P12.5), the P12 category represents a dependent variable in the decision, rather than an independent one. This is not to say that it is unimportant. It may well represent the way in which instructors justify the decision to themselves, or conclude the process. But since it is clearly dependent upon other decision factors being resolved, it will be placed last in the positive process of the solo decision model (Figure 9.8), represented as the instructors' reflection on the reasons for allowing the flight (the final conclusion prior to the solo outcome).

There is one exception to this. Sub-category P12.5 ('extra capacity') does appear to provide an independent decision attribute. A number of instructors pointed out that they were inclined to send students solo because as well as flying safely, the student 'talked through' what they were doing 'out loud'. According to several instructors, this demonstrated that the student had "spare capacity". For example:

"When it's fairly busy on the airfield... and they're talking to you. They're saying, 'um, a bit tight on the aerotow, not sure what this cable's going to do, I'm going to land long'"

No instructors mentioned 'not talking' as a reason to disallow solo flight, and hence this appears to be a factor that loads onto the positive decision outcome only, and hence appears to used by instructors to confirm that a student can safely fly solo. This was represented as a separate category for the purposes of the decision model, and was therefore grouped with all other independent 'positive' categories.

All positive independent decision attributes in the model are grouped, since combining rules are not known (see Figure 9.8). These were categories 'P/N' 2, 4, 7, 10 and 12.5).

Consistency (N8 and P8)

In its positive form, 'consistency' appeared to be a way in which a lack of negative occurrences was articulated by participants as loading onto the positive decision outcome. Hence, like P12, it was often used more as an expression of combining rules than as an individual factor affecting the decision. This would further support the non-compensatory notion that no problematic events (i.e. causal factors) must occur if the student is to fly solo. An example was:

"consistently landing well, is a big thing, consistency in landing"

From this comment, it can be seen that consistency was not an attribute independent of other attributes but a vehicle used to express longitudinal assessment by the instructor, and as such a 'combining rule'.

From a decision making perspective however this application of the term '*consistency*' appears redundant in the present context. In the above quote (as well as many others), *consistency* refers to a lack of causal events (bad landings), and in the same way ‘*inconsistency*’ relates to one or more causal events (bad landings) amongst satisfactory ones. For example

“..they couldn’t consistently get the aircraft down safely.. inconsistency in her ability to flair and land the aircraft”

It can be seen that “*inconsistency*” as used in the above quote (and many others) is a factor of the combined ‘safe’ and ‘unsafe’ landings that the pilot had made, but not a feature of any of those landings. Since the ‘safe’ landings were clearly not the cause of the instructor preventing solo, these could be dropped from the decision analysis, which would leave only unsafe landings (at least one). But unsafe landings are already catered for in the decision process. Hence to use of the term ‘consistency’ seems unnecessary, and appears to be redundant in that all the assessment of necessary decision attributes must have been satisfied without the need for it, whether positive or negative. Hence it appears to be a way of instructors articulating the decision process or their overall analysis, rather than a real consideration. Some comments in the ‘negative’ category were slightly different and demonstrated the use of *consistency* to illuminate a combining rule. A typical example was:

“making the same bad decision twice, on two flights, having previously talked about it..”

In this comment (and others) the instructor’s decision was affected by the same thing happening twice. It is possible that the decision error may have been below the threshold ‘cut-off’ point of a non-compensatory negative factor (and hence not prevented solo flying had it occurred once only), but the reoccurrence of it meant that it did prevent solo flight. This hints at the possibility that specific causal factors, not in themselves sufficient to prevent solo flight, can combine to reach that threshold (albeit still in a non-compensatory manner, since one factor cannot reduce another, only add to it). This may be because of what the instructor infers from these happenings, such as lack of learning on the part of the student, or overall inability in one area. Other comments in this category hinted at a further possibility, and suggested that numbers of otherwise small errors in different areas could add up to prevent solo flight, possibly without any one 'cut-off' point being reached. For example:

"I think, all the way round the circuit there were things to criticise there"

Hence it is possible that some of the specific 'non-compensatory' factors do operate together when numbers of less serious events are combined. Because this is the only use of 'consistency' that was not already accounted for by other factors, it was included in the decision model as a method of combining attributes to cause a negative decision outcome (no solo). The present research, while recognising this possibility, has no data suitable for further investigating it. Importantly there is still no suggestion, or evidence, that combinations of 'causal-factors' use any form of compensatory combining rules. The suggestion is simply that some additive rules may apply, but no deductive rules. Importantly there is no evidence that one factor can compensate for the occurrence of another.

Lack of consistency (N8) was represented in the decision model (Figure 9.8) as a component that offers an alternative method of combining the six non-compensatory attributes to prevent solo flight. As opposed to a single attribute simply reaching its cut-off point, this highlighted the longitudinal assessment that allowed 'low-level' causal events to combine to a threshold sufficient to prevent the solo. The exact dynamics cannot be known, and would require further research.

The reason it did not load onto the positive outcome is because the way it was used by participants was redundant to P12. For instance, if a pilot consistently made good landings then they made no bad ones, and hence this was accounted for by the combining rules applied to the six 'specific' non-compensatory attributes.

In the model, no arrow leads into the 'performance assessment' category, simply because unlike individual causal factors, there was no evidence that multiple low-level factors could be compensated for by performance assessment. In reality this may well occur however, and so more research is required on this point.

Indirect Issues (N7 and P7)

Many critical incidents in this category related to instructors' impressions of their students such as "confident", "relaxed", "methodological" and "enthusiastic", as well as the important aspect of 'self awareness' (in terms of performance assessment). These appear to be important independent parts of the decision process, and are reflected by smaller parallel negative categories. As well as being used to confirm instructor judgement regarding solo flight, 'indirect' issues were also used by instructors to prevent students being allowed to fly solo. So despite no causal-factor incidents occurring, some students were prevented from soloing on the basis of some indirect issues, as exemplified by reasonably large numbers of comments such as:

"so bloody good, but so cocky with it.. I'm in command of this thing, this is a doddle, it's like playing on the computer". But it isn't"

"It wasn't so much that he hadn't been doing things right , but he hadn't known that he hadn't been doing things right... the self diagnosis type thing..so I think that to me ...that says 'I'm not ready, or you're not ready'.."

This supports the notion that although solo is often disallowed on the basis of specific performance issues, more than the absence of such issues is required to allow solo flight. The pilots in above quotes made no errors or mistakes and hence were not discounted from solo flight in the initial 'non-compensatory' part of the process (shown by Figure 9.7).The decision model reflects this with arrows towards both 'solo' and 'no solo' outcomes from the 'positive' group. 'Performance assessment' (as seen in the second quote) has been discussed as the only obvious factor that positively compensates for specific causal events. However it also appears to function as a further decision factor that can add to an instructors' positive decision in general.

A further interesting (although small) category concerned pressures felt by the instructor. This was exclusive to positive decision outcomes, and suggests that some instructors are influenced in their decision making by factors not related to their consideration of the safety of the student to fly alone. However, being a small category, there is no evidence on which to further hypothesise this point.

From the incidents and themes emerging from the ‘indirect categories’, there was sufficient evidence to hypothesis that indirect issues were considered as well as the non-compensatory (technical) parts of the decision.

Instant reactions / behind the glider (P2 and N2)

These were small categories, containing just 13 incidents between them. Because of this, little can be deduced. However the theme of the comments was consistent and appeared to be an independent decision attribute. Examples are:

“Instant reaction to the sink that you get over the Tring road... flying along and there’s this sshhhoo the nose goes down, turns in.. Just the reaction and the erm, the response to the unusual, that’s very good.”(P2)

“He just wasn’t on top of it, you know, he was behind the aircraft... a toboggan run, you know, a runaway train” (N2)

“They did not react at the rate that I would have wanted them to react”

It can be seen that the instructors did not mention anything specifically going wrong (no causal accident factor) despite an overall assessment. However in the positive comments instructors were more specific. This could represent a heuristic device that instructors use for helping to determine the overall skill level of students, but further research and more data would be required to establish this.

Sufficient evidence is available to determine that these categories are relatively independent decision dimensions, and so they were included in the group with other such categories in the decision model.

Situation - Instructor harmony / disharmony (P4 and N4)

P4 is perhaps the positive category of most importance, in terms of both concept and critical incident numbers, with three times as many incidents as N4. It refers to 'non-tangible' concepts that instructors often talked about when making positive decisions. Common expressions were 'gut feeling' (P4.1.2), 'confidence in the student' (P4.1.1) and 'feeling comfortable' (4.2.3) to name a few. These concepts were hardly mentioned with regards to disallowing solo flight. Many similar sub-categories arose from negative comments, but not to the same degree as with positive comments. Indeed most of the negative categories of this sort nearly disappeared in the participation rate check.

This category therefore lends considerable support to the notion that when sending pilots solo instructors need evidence over and above an observably safe flight, evidence that they do not commonly need to rely on when disallowing solo flying. The non-tangible nature of comments in this category demonstrates that instructors often find it difficult to declare knowledge relevant to sending pilots solo. This was stated by a number of instructors in the initial pilot interviews prior to the CIT research.

Many instructors mentioned that the situation had been different from what seemed right to them, without being able to describe what was different about it. Also there were numerous similar 'positive' comments where the instructor had difficulty in expressing what they 'felt' was right about the student performance. This indicates the use of non-declarative knowledge in decision making and the possibility of an intuitive element to the decision process (as described by Klein 1998). This could occur when a well-known situation either looked 'typical' or had a slightly different pattern to what looked familiar (e.g. the pilot not looking in the direction that pilots normally look during a particular part of the flight). In these situations an instructor may not be aware of the exact nature of the cues that drive their judgement, since they are the result of the activation (or not) of contextual schemata (see Reason 1990). Due to the subtlety of these sorts of patterns it is common for people to be unable to describe what they have observed (Klein 1998).

Because instructors may have been quite certain that something was right (or wrong) without knowing what it was, they tended to use terms such as ‘gut feeling’, and ‘feeling uncomfortable’. A number of comments made it clear that, even from a negative decision perspective, this process is not the same as the non-compensatory process described early, for example;

“*You know that feeling where you just want to take over all the way round, but you don’t. It was like that, it never really felt right.*”

The further context around this comment (and others of this sort) shows that the non-compensatory process returned no definite ‘specific’ problems but that the overall ‘intuitive’ decision over-rode this to disallow solo flight. Other comments made it clear that the intuitive process had favoured solo flight, but only when there were no observable causal factors. Importantly, no comments suggested that a specific ‘causal’ factor could be positively compensated for by overall intuition. Comments relating to intuitive processes of this sort were not used in conjunction with specific observable causal events. This further supports the non-compensatory nature of the ‘causal factor’ critical incident categories. It appears that instructors do not need to rely on intuitive processes if a student makes an overt error, since the decision outcome is then clear.

More than any other category, P4 and N4 represent non-declarative knowledge leading to ‘expert judgement’ or ‘intuition’. The critical incident numbers suggest that either instructors rely on this more when making positive decisions, or require it less when making negative ones, perhaps because the non-compensatory process often eliminates the need (Figure 9.7).

Due to the nature of expert intuition, it is problematic to determine the extent to which the outcome is dependent upon other attributes in the decision. For these reasons, in terms of the decision model, P4 and N4 were grouped along with all other 'independent' positively dominant categories.

Dealing with Emergencies / not coping (P10 and N10)

Literature advises that pilots must be able to deal with emergencies prior to solo, for example "your instructor, when sending you solo… considers you competent to handle an emergency" (Thom 1997) and "prove that you can.. deal with any contingencies such as cable breaks, running short of height, and stalling and spinning" Piggott 1997), to quote just a few examples.

Some comments in N10 (not coping) shared many similarities with N4 (Situation - Instructor disharmony) as being related to intuitive processes, but expressed this as concern that the student will not cope in the future, particularly with an emergency. For example:

"Some sort of third instinct that would say if he had a cable break he may not do the right thing. He might decide to land ahead when he couldn't, or he might decide to turn when he should have landed ahead"

Other comments were related to symptoms that students showed during the session, and as such these could be quite specific, for example

"he just froze on the controls. He was flying at great speed towards the hill and would not do anything, wouldn't even let go of the controls... very tense".

Hence, whether or not these were intuitively based or directly observable, they represent a further possible independent decision attribute for the instructor to consider. A greater proportion of positive incidents were categorised in P10 than negative incidents in N10. In a positive sense this category was concerned only with potential emergencies (simulated and real). In common with the negative comments, some positive comments were based on intuition, others on specific observable issues. For example:

"I threw the book at this bloke... he coped, difficult cable breaks, actually a very gradual power failure... he just ripped through it all"

Clearly this could be used by an instructor to support trouble free flying and allow the pilot to fly solo.

P10 and N10 were therefore represented in the group along with the other independent attributes in the decision model. Since most of these attributes were both positive and negative (despite being 'positively dominant'), arrows led from this group towards 'solo' and 'no solo' directly.

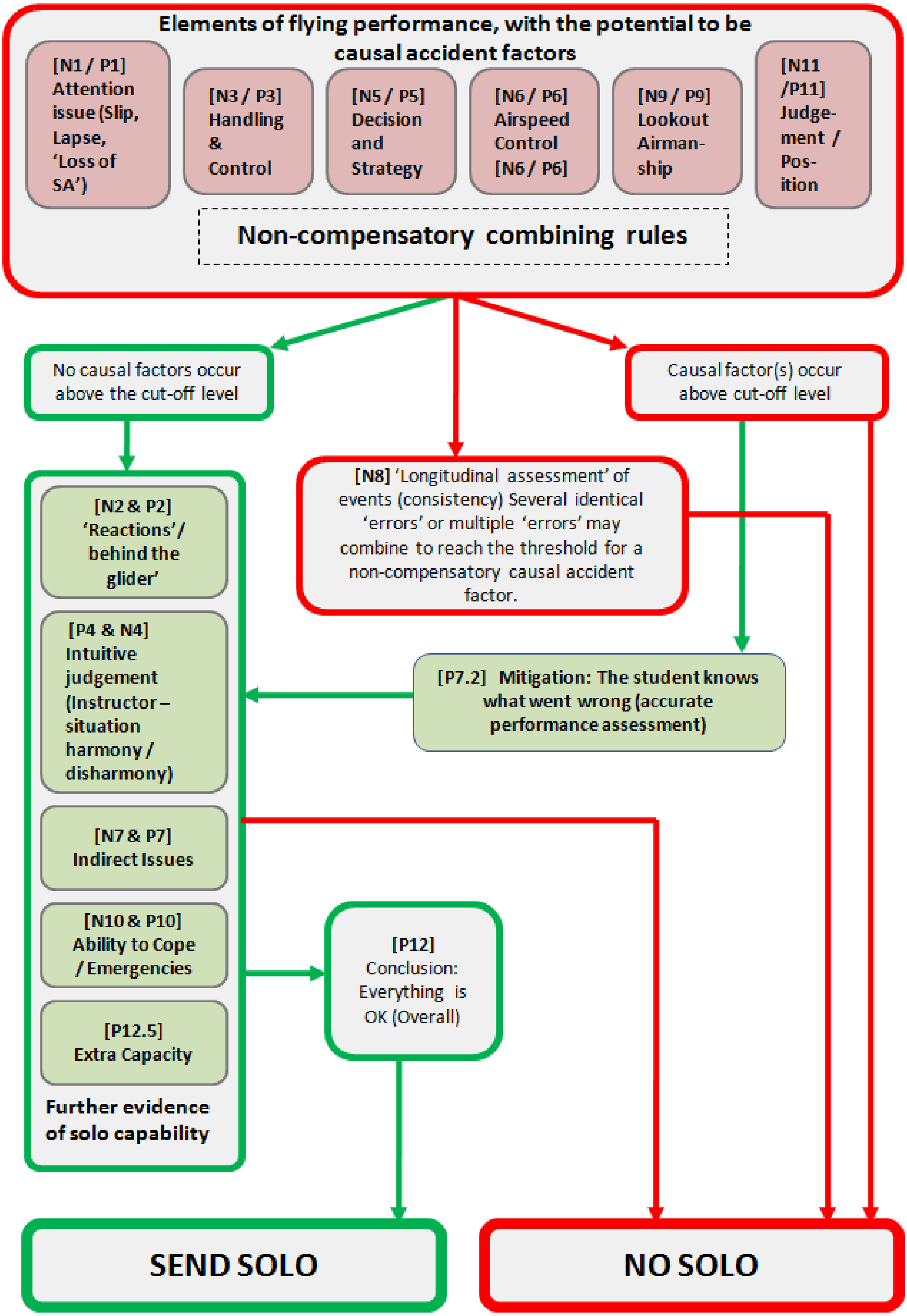


Figure 9.8 Model of instructor decision making, from CIT data analysis

The decision model

The final model is shown as Figure 9.8, and flows from top to bottom. Each of the 23 top-level categories are represented (and labelled) with the exception of P8 (consistency) because there was no evidence to show that it was anything other than a redundant expression used by instructors, as explained). Additionally P7.2 (‘performance assessment’) is shown as an independent attribute as previously explained. Each category (or in many cases ‘category pair’) represents a decision attribute. Tentative links between them have been added where evidence from incidents warranted it, to show the most logical way in which they fit together.

Fundamentally, the model follows two processes to arrive at a positive solo decision.

1. Were there any observable ‘specific’ events with the potential to cause an accident? (this is represented by the ‘negative’ group of categories)

2. If not, is there sufficient further justification (further confirmation) from other sources that can confirm that the student should go solo? (This is the positive group in the model, e.g. ‘pilot attitude’, ‘extra capacity’, ‘reaction times’, ‘instructor intuition’).

The model basically consists of a negative group and a positive group, reflecting the two parts of the process. Although combining rules are hypothesised for the ‘negative’ group, there was no evidence with which to do so for the ‘positive’ group.

The model builds on the previous model of the non-compensatory process (Figure 9.7). The majority of these links simply relate the category to the appropriate decision outcome; ‘so solo’ or ‘solo’ (via P12). These could be confidently ‘mapped’ because each negative critical incident (making up most of each ‘negative’ category) led to solo flight being disallowed, and each positive one led to solo flight being allowed (by definition). This is the basis upon which the direction of the majority of linking arrows has been added, and is therefore easily justifiable. The ‘causal factor’ type categories are coloured red to denote that they are predominantly negative categories, and green for the ‘positive ones. Parts of the decision process such as those indicating combining rules (N8, ‘consistency - negative’ and P12 ‘general, overall’) are ‘outlined’ rather than ‘filled’ to indicate that they are not considered decision attributes in themselves.

Red and green arrowed lines only seek to show logical interpretations of the way in which categories combine, and not to attempt to map the process itself, since this would require more research. However all arrows are supported by evidence from the critical incident data. For example no arrows exist between ‘longitudinal assessment’ (N8) and the positive decision side of the model, since no critical incidents indicted such a link. This includes the mitigating effect of accurate self assessment. No instructors claimed to have sent a student solo on this basis, but this may need further exploration in future. Arrows are red if they go towards a ‘negative’ attribute and green if they point towards a positive one. In this way the positive and negative elements of the decision are made clearer.

There is direct evidence from the data for the existence of all the elements in the model, most of which are categories containing data from the instructors. The critical incidents show that all of the factors have been involved in the decisions to send pilots solo or prevent them flying solo. Hence the categories represent the attributes of the decision.

However the model is not intended to represent a decision algorithm, although where evidence was available this has been tentatively hypothesised (i.e. non-compensatory combining rules). Although evidence exists for each element, and for some of the dependencies within the decision, the exact nature of the decision process cannot be known from the data available. This includes the exact nature of combining rules and the weighting of elements within the model.

Conclusion (Section 1)

It has been shown that the two dichotomous decision options (allowing and not allowing solo flight) are not simply complimentary processes, i.e. if there is insufficient evidence to prevent solo flight it does not mean that there is sufficient evidence to allow it.

The evidence suggests that problem-free flight is insufficient for many instructors to allow solo flight. After observing any number of safe flights, the instructor looks for further confirmatory (positive) evidence to 'sure up' a decision to allow solo flight ('pilot attitude', 'good self critique' etc). It is well known that "people show a strong tendency to accumulate several instances of confirming evidence" (Hogarth 1987).

This overall view is supported by instructor comments showing clear decision conflict. This happened when the instructor observed no causal accident factors (good) but then found no further 'positive' evidence to build a positive decision outcome upon (such as 'a good de-brief', 'a feeling of confidence' or evidence of 'extra capacity').

This research demonstrates the advantage of an exploratory approach when the structure of the decision is not fully understood. The findings were very similar to those of Ikomi and Guion (2000) in terms of the main reasons why students were not sent solo (non-compensatory combining rules applied to specific elements of flight), but showed a further aspect to the decision making process particularly related to choosing when a student can go solo. Ikomi and Guion (2000) found unexpected leniency on the part of the instructors to send pilots solo (on paper). It is possible that this was caused by the experiment only accounting for one part of the decision process (i.e. the non-compensatory part). The present study shows that the processes of sending a pilot solo and not sending them solo are not complimentary (as would be logical). This shows the benefit of using an exploratory technique before modelling a decision process based on assumptions.

The method employed by the present study makes it likely that all decision elements have now been identified. The data were shown to be exhaustive of the subject area by robust processes (see Figure 9.1), as well numerous validity checks and SME reviews. Therefore, because the decision model is based exclusively on the data, it can be confidently claimed to be exhaustive in terms of decision elements (within the limitations of the interview method).

Importantly however, whereas the model is probably representative of all the attributes driving the decision process, there is no evidence that any instructors use all of those elements in coming to a decision, and hence no evidence that any instructor goes through the process in its entirety. This is a critical point. No instructors mentioned all elements (no element had a 100% participation rate, the highest was 82%, for N11 [Judgement and Positioning]). Whereas this does not provide strong evidence against the holistic nature of the model, it nevertheless leaves the possibility that certain instructors apply certain parts of this process only. This is an important area for future research, along with further field validation of the model and investigation into the dynamics of the decision mechanism.

The parts of the model containing the 'positive' (and general) attributes require a great deal more research. It is not known how these attributes link to the intuitive decision processes (N4 and P4), or indeed whether they are elements of the same process. Although all the elements of this group are accounted for within the critical incident data, no attempt was made to infer combining rules, dynamics or weightings. The decision model can only be used to say that these attributes are (and have been) used by instructors when considering solo decisions, and that it is likely that instructors use them to provide additional evidence (above the student technical ability) to help make 'right' decision.

The main value of this analysis is in exposing the elements of what is possibly the most safety critical decision made about a pilot in his or her lifetime, and opening many doors to future research of a more practical nature. The dynamics of the decision require more mapping. More needs to be done to look at how specific flight elements trigger a no-solo decision, but more importantly which elements are most likely not to trigger the mechanism. The manner in which evidence is weighed between the components also demands attention, because the method of evaluating positive attributes could lend itself to biases such as confirmation bias (Wason 1960). The use of intuitive processes may or may not lead to premature solos and more needs to be done in order to understand these processes.

The most pressing piece of research is a field validation of the model, and evaluation of the extent to which instructors use it in a holistic sense.

Section 2: Comparison of critical incidents with previous accident analysis

The critical incident findings were directly compared to findings relating to flight phase and accident cause for newly-soloed pilots, from previous studies (chapters 5 and 6).

Only critical incidents of a sufficiently specific nature (e.g. in-flight events) were categorised by flight phase, using the six main flight phases from Jarvis and Harris (2008). Notably, 82% of these led to 'negative' decision outcomes. Almost all the same incidents were found to belong to the 'causal factor' categories which aligned with the accident causations from chapter six (attention, judgement, handling and strategy). Hence, when comparing all the critical incidents to findings from accident data, many did not apply. Only those in the 'causal factor' categories (N/P1, 3, 5, 6, 9 and 11) or coded as flight phases one to six, were looked at. This made up about 60% of critical incidents. It was hypothesised that these represented instructor observations that were sufficient to disallow solo flight, and could not usually be compensated for in other ways (see Figure 9.8).

Comparison of proportions of causal accident events (from the CIT data) with the four top-level accident categories (from chapter 6) showed broad alignment. There were similar proportions of the 'attentional' and 'decision-making' issues within the two data sets (Table 9.3). Together, these made up only a small total proportion of both critical incidents and accident events to low hours pilots (less than a quarter). This is because training flights and early solo flights are generally confined to local area flying, and have fewer opportunities for strategic decision mistakes (Jarvis & Harris 2008). Additionally, many accident causes categorised as 'attentional' occurred with more advanced glider types, being concerned with control slips and lapses relating to undercarriage and flaps (see chapter 6). Hence these were not applicable to low hours pilots. Training gliders are more basic, with fewer controls to forget or confuse. Highly inexperienced glider pilots are restricted to flying training gliders or basic single-seat glider types.

Handling and judgement categories accounted for the vast majority of accidents and causal critical incidents. Proportionally more accident events (for early-solo pilots) related to judgement (53.6%) than handling (32%). This trend was reversed for the critical incidents (42.6% and 35.8% respectively). This is probably because almost all judgement related causal accident events (for low-hours pilots) occurred in the approach and landing phases (Table 9.3), but these phases were far less represented in the CIT data than the accident data. It is also possible that handling issues are easier for instructors to identify and comment upon than judgement issues. An instructor can directly observe handling issues by shadowing a student's control movements and noticing a number of clear and instant indications (yaw string, airspeed indicator and horizon). However instructors cannot directly observe student judgement.

In terms of flight phase data, the two most striking results were those for the circuit and approach phases, but the launch and landing phases were also of some interest. The following section discusses the results for the circuit first, because this is possibly the area that instructors perform best in terms of assessing student problems. The launch and landing are discussed next and the approach phase last because it demands most attention.

The Circuit

Figure 9.5 shows that over half of the 'specific' critical incidents were related to the circuit phase. No other phase accounted for more than 15% of 'specific' critical incidents, and so the circuit phase represented a considerable departure from the norm. This indicates that instructors base non-solo decisions on the students' performance in the circuit much more than on other flight phases. This phase was not unusual in terms of proportions of positive and negative comments (85% negative).

Instructors reported causal events of all kinds taking place in the circuit, and Table 9.8 shows that the majority were involved with the positioning of the glider relative to the landing area (circuit judgement). This is in line with gliding literature which highlights the importance of judgement to achieve the right circuit position due to the nonpredictability of this phase compared to powered forms of aviation (Piggott 1991, Stewart 1994).

It is possible that the high number of critical incidents reflects a higher number of such events taking place in the circuit, but accident data do not necessarily support such a hypothesis. Only 11% of accidents to early solo pilots were caused in the circuit (Jarvis & Harris 2008) which would suggest that causal-factors in the circuit are not as common as in other phases. However it is also possible that the disproportionate number of critical incidents relating to the circuit helps to explain this lower accident total for early solo pilots. Such accidents may be fewer due to more effective 'filtering' of these problems prior to solo, explaining the high numbers of negative critical incidents relating to this phase. Therefore high numbers of critical incidents and a low accident total for inexperienced pilots could be a sign of instructor effectiveness in terms of training and assessing circuit flying.

In terms of causal factors, critical incidents relating to the circuit spawned 28 low-level subcategories, meaning that instructors observed numerous circuit problems that affected their decision outcomes, mainly to do with judgement and handling. None of the 28 categories were particularly large, indicating that there was a wide spread of issues. Other flight phases had fewer and 'stronger' emerging themes, which provides a further explanation to the proliferation of critical incidents related to circuit flying. The circuit takes up far more of the flight than the phases that cause most accidents (approach and landing). Arguably it also has more components and complexities. This means that instructors have more opportunity to notice and consider minor errors, and students have more opportunity to make them. Additionally it is possible that the instructors' workload is much lower in the circuit than the launch, approach and landing, due to the extra time available and a much bigger margin for student error. These factors could explain the breadth of circuit related issues raised by instructors, and the possibility that errors in the circuit are more likely (than other phases) to be noticed.

Because there was no way of establishing the weighting of decision factors, it cannot be not known how critical the various circuit issues were to instructor decision-making compared to those of other flight phases.

The number of critical incidents related to the circuit is disproportional to the other phases, as well as being out of step with the relative number of circuit related accidents to low hours pilots. More research is required to uncover whether the amount of attention instructors give to this phase of flight is effective in maintaining safety. It is possible that it has a protective effect, helping to maintain a low accident rate in early solo pilots. Hence the current situation requires further research before any attempt is made to redress the balance of the decision attributes in terms of flight phases.

The Landing

Early solo pilots have a high rate of landing accidents (Jarvis & Harris 2008). Figure 9.6 shows that present results for the landing phase align with previous studies (Jarvis & Harris 2008, and Chapter 8). As well as being the second most common accident cause to early hours pilots, it had the second highest number of related critical incidents and was rated by instructors as the second most likely phase to cause accidents to early-solo pilots (see chapter 8). The causes given by instructors for not sending pilots solo (critical incidents) were also in harmony with the findings from accident data from chapter six ('judgement' and 'handling', almost exclusively in the round out).

The launch

Launch accidents were those perceived as most likely by the instructor to occur to low hours pilots (see Figure 9.6), and possibly those most feared (see chapter 8). This was reflected by a larger proportion of critical incidents than accidents to early solo pilots (12% and 7% respectively), showing possible caution on the part of instructors with respect to this phase. It was notable that of 42 critical incidents, 41 refer to 'handling' issues. This is in line with the fact that the rotation into the launch climb was associated with most fatalities from 2002 - 2006 (Jarvis & Harris 2008), and supported by discussion from chapter eight concerning BGA publicity around serious winch launch accidents of this type in 2007. These serious accidents were almost certainly issues of glider handling (applying too much backward pressure to the control column, or allowing the glider to pitch up too rapidly, depending on the glider type). Hence it appears that instructors show some caution with respect to the way pilots handle the launch phase of flight. No such serious accidents occurred to early-solo pilots from 2002 - 2006. Although this was not statistically likely due to the small numbers involved (see chapter 4 Table 4.3, and Chapter 6), it could be that instructor sensitivity to this area has helped prevent such accidents.

The Approach

Whereas the circuit is associated with a high number of critical incidents and a low number of early solo accidents, the approach phase shows the opposite trend.

The approach is associated with the highest accident rate of all phases for low-hours pilots (causing 39% of accidents) and is 10 times higher than for more experienced pilots (Jarvis & Harris 2008). Against this background, there was an expectation of finding numerous critical incidents to provide further explanation of this problem. However, only 20 critical incidents out of 659 related to the approach, 17 of which could be categorised into a causal factor category.

It was previously found that all accidents initiated during approach (to low-hours pilots) were caused by judgement and handling issues (see chapter 6). Additionally all fell into just three relatively small categories; undershooting the landing area (caused by misjudgement of the approach path leading to too much airbrake being sustained), mishandling the glider in pitch, and having too little airspeed). It is therefore of considerable note that of the 17 critical incidents none were categorised as relating to judgement. Furthermore only five were related to handling, the remaining 12 being associated with strategy / decision-making (Table 9.8). The findings from chapter six showed that no such accidents happened to early solo pilots during the approach.

The majority of the ‘decision making’ descriptions were due to two sources. A number were concerned with the decision to land straight ahead after a launch failure (which is a unique type of manoeuvre). This was treated as a separate sub-phase by Jarvis and Harris (2008) and contained no accidents to low-hours pilots. A further seven were descriptions of problems relating to the direction of the approach (laterally), such as approaching towards another glider or vehicle. There were no such accidents to low-hours pilots from 2002 - 2006 (see chapter 6).

Of the remaining five critical incidents, two were associated with ‘positive’ outcomes (hence reasons why pilots went solo). This left just three critical incidents that related in any way to the reasons for accidents identified in chapter 6. Hence the critical incidents were out of step with the accident findings, in terms of total numbers as well as the reasons given for the ‘solo decision’.

As was suggested with the circuit phase, the contrast in the number of critical incidents to accidents could help explain the accident figures. If specific problems are not being picked up during training (which seems likely given the analysis of critical incidents) then it helps to explain why those factors cause such a high accident rate for newly soloed pilots.

It was previously argued that the same issues that caused accidents to low hours pilots should be numerous and apparent during instructional flights prior to solo (Heinrich’s accident and incident ratio, Heinrich 1980). The critical incident data suggest that either this is not the case, or that if it is, the problems are not noticed, not remembered, or not deemed serious enough by instructors to disallow solo flight.

Accidents caused in the approach phase rank third in frequency during instructional flights (2002 - 2006) behind launch and landing Table 5.10 Chapter five. This strongly suggests that causal accident factors do occur on the approach, while students are still under instruction.

The finding that the approach phase was ranked fifth by instructors in terms of likelihood of causing accidents (Chapter 8) supports the present finding in general and shows that instructors do not believe that the approach phase presents the greatest risk of an accident to early solo pilots. If instructors very rarely prevent solo flights due to issues occurring during the approach, then such incidents would have lower 'availability' in terms of recall than other events (simply based on frequency of episodes).

Unfortunately little explanation can be gained from the analysis of critical incidents relating to the approach due to the small number of appropriate cases. The BGA instructors' manual (British Gliding Association 2003) includes a comprehensive section on approach control including a large variety of exercises to show problems that can occur and how to deal with them (including undershooting, and the effect of not closing the brakes to recover from it). Most of these are written as instructor demonstrations however (rather than student exercises), and the only exercise related to checking the student for 'undershoot tendencies' involves the instructor setting up a low final turn to check that the trainee does not open the airbrakes immediately (termed 'landing lever syndrome'). There are no exercises that involve the student demonstrating that they can recognize or recover properly from an undershoot situation, which was the essence of about half of the accidents to early solo pilots (see chapter 6). Given the accident profile of early-solo pilots (See results of chapter 5 and chapter 6) this appears to be a serious omission, particularly as advice to instructors in the same document (in the section about first solos) is to check that students are able to recognize and correct for an undershoot prior to solo (British Gliding Association 2003).

Clearly, without setting up such a situation, the only way that instructors can check students' recognition of an undershoot is if one happens for real. It is possible that the findings of minimal critical incidents and low instructor concern are explained by a lack of real undershoots occurring on instructional flights (combined with no simulated ones) and hence no opportunity for the student to misjudge them. If the first opportunity a student gets is when they are solo, then the accident rate is not difficult to explain. There are numerous hypothetical reasons why students are less likely to get into such situations whilst under instruction. These include briefings (e.g. reminders of wind strength), prompting and instructors taking over prior to a situation conducive to a problem presenting itself (e.g. a low final turn).

However, even if students did get the opportunity to "recognize and correct for an undershoot" during training, it is questionable whether an instructor would be able to recognize the problem. Considering the length and size of most gliding airfields (required for launching gliders) it is possibility that many pre-solo undershoots go undetected by instructors because they have no way of telling exactly where the student was trying to land on the field (most gliding sites are large featureless grass airfields). Since gliders cannot taxi and have no power, there are no designated landing points as there are in other forms of aviation, because other gliders may be occupying those areas at the time. Hence gliders tend to land wherever is most safe or convenient. This is why chapter 12 of the BGA instructors' manual (approach control) refers to the need for the instructor and student to agree a reference point (RP) before doing approach control exercises ("agree an RP, perhaps in relation to a parked car or glider"). This is because approach technique in gliders relies on judging the approach angle of the glider relative to a reference point directly ahead (just prior to the touchdown point). Without agreeing such a point, the instructor cannot know where the student is trying to land. The instructors' manual contains much advice on interpreting the relative movement of the RP during approach, but since the advice is to pick an RP "relative to another object" it is not clear how this can be achieved. Furthermore this advice is only given for specific approach control demonstrations, not for all flights. The very fact that the BGA deem this necessary for demonstrating approach control to students supports the proclamation that a problem exists during normal flights (i.e. where an unexpected undershoot may develop). This helps to explain why approach judgment was not mentioned in any critical incidents at all. Without specifically setting up the situation there is no way for an instructor to tell if the student has misjudged the approach path or not, unless specific circumstances combine with a very serious undershoot. As has been discussed, such circumstances are probably less likely to occur while pilots are under instruction.

Added to this problem, the BGA instructors' manual advises instructors that students need not achieve spot landings in order to go solo (British Gliding Association 2003). Indeed this sentiment was mentioned in some 'positive' critical incidents, for example:

"Not a spot landing, but a good landing. I don't care if it's a spot landing as long as it's safely on the airfield"

Given the high accident rate due to misjudgments of the approach path, this could be seriously ill-conceived advice. If a student does not plan to land at a particular point, then they deprive themselves of feedback in terms of their technique ('knowledge of performance') and the outcome ('knowledge of results'). Both these forms of feedback are crucial for development of perceptual motor skills (Magill 1989). Additionally, the instructor cannot tell if the student has achieved what they intended and hence cannot gauge the approach performance or give further feedback. In many situations this means that it is quite possible for student pilot to simply accept whatever approach path the airbrake setting provides, particularly on large airfields, and hence not practice the skills and judgment required.

Once alone in a glider, this would be a latent training failure that would perhaps be exposed the first time the pilot tries to land a little nearer the airfield boundary, but perhaps linger on much later into the pilots career when they must approach accurately into the upwind end of a small out-field. Accident statistics show that outfield landings are by far the largest source of accidents in gliding, and although many start as decision making accidents, they often end with undershoots and overshoots. Hence the issue of how students are sent solo could have much wider implications for the lifetime of the pilot.

The finding that so few critical incidents related to the approach, and fewer still to the most critical aspects, is itself extremely instructive considering it is the major area of accident causation for early hours glider pilots (Jarvis & Harris 2008). This makes the approach phase an even more notable omission in the work of Ikomi and Guion (2000) since it would have made a very interesting comparison to the present results.

Further research is required into the method of teaching and assessing approach control and judgement in gliders. The system may require serious overhaul. It is possible that if some form of reference point is marked on airfields (e.g. a painted line perpendicular to the landing area) then this would enable feedback for the instructor and the trainee during training. However further research is needed into this area.

Conclusion - Section 2

The results suggest the possibility that instructors properly assess those aspects of the launch, circuit and landing that are most likely to be hazardous to student pilots once solo. A low level of accidents to low-hours pilots, particularly in the circuit and launch phase, may be partly due to appropriate attention paid by instructors while making 'solo decisions'. Although there is a high rate of accidents in the landing phase, the research suggests that instructors are aware of this, and also that they regularly prevent solo flights due to the occurrences on landing that are similar to those causing the high rate of accidents to low hours pilots. Hence the instructors are not out of step with landing issues, and may indeed be responsible for preventing more accidents in this phase.

However the approach phase shows a very different pattern. Approach path judgement, speed and pitch handling caused all approach accidents to early hours pilots from 2002 - 2006. However instructors appear to be unaware of the level of accidents caused to early solo pilots on the approach (see Chapter 8). Furthermore the data suggests that instructors rarely (if ever) prevent solo flying on the basis of judgement or handling issues in this phase, despite these causing the highest accident rate to this already vulnerable group of pilots. Based on the results and the gliding literature, the reason behind this is that during training flights there is less chance of circumstances arising in which these events occur. Furthermore when such events do occur, instructors fail to notice due to a lack of cues available. This could mean that pilots do not properly learn safe approach control until after they are solo, which could account for the results of previous studies (Jarvis and Harris 2007b, 2008).

Piggott could be more accurate than perhaps intended when he claimed that “Approach is a busy time for the beginner and it is only some time after soloing that he learns to organize the thinking and flying so that there is enough time for refinements” (Piggott 1997). Unfortunately it appears that this learning time may account for more than just refinements, and while it is occurring the pilot is particularly vulnerable to an accident.

Overall Conclusion (Chapter 9)

The overall aims of this study were to gain an initial understanding of the decision making process of instructors with regard to allowing solo flight and to explore in what ways instructor decisions could be contributing to the high accident rate of newly soloed glider pilots. A decision model was created, based on findings from critical incident data that highlighted the complexity of the decision process, as well as the numerous elements involved.

A fundamental part of this process is the way in which potential causal accident events (observed by instructors) discount from solo flight in a non-compensatory manner. Hypothetically, if the components of this part of the process were aligned accurately with accident causes, then it would act as a filter to potential accidents. However this is reliant on the instructor noticing these potential accident events and assessing them with appropriate weighting. It was found that this process may work for most parts of the flight, but is certainly ineffective in terms of the approach phase and hence may be partly accountable for the high rate of accidents to new solo pilots caused in this phase.

Further research is needed to validate the decision model in the naturalistic environment. This would help to assess the dynamics of the model which are tentative, and the weightings given to the various components, which are unknown.

It is recommended that the training and assessment of approach control be reviewed in light of the present findings, and instructors made aware of the high accident rates caused by this phase, and the reasons behind them.

Chapter 10 - Conclusions and Recommendations

Much has been discovered over the course of these studies, and the implications are far reaching in the gliding community, as well as having application to other areas of aviation and research. Conclusions and recommendations were given with each piece of research. Rather than repeat these at length, this section summarises the main conclusions and recommendations.

Prior to the commencement of the work, no knowledge existed concerning the accident profile of glider pilots. It was identified that early-solo pilots had a different accident profile to more experienced pilots. Popular notions that early solo pilots were safest were shown to be mistaken. The issue of pilot judgement in approach and landing was uncovered as a major factor in most accidents to low-hours pilots. Instructors were found to have the same misconceptions as the popular literature in the terms of the relative safety of these pilots.

It was found that instructor decision making in terms of solo flight decisions was out of line with the types of accidents that students were most likely to have when sent solo. The project shows this decision to be a highly complex problem for instructors, involving numerous decision factors.

As well as investigating these specific issues, this work challenged a number of methodologies. These centre around the use of accidents as populations rather than samples of all flights. Accident counts are often mistakenly used in this way (as shown in chapters one to three). By assuming that accidents are a population, false conclusions can be drawn regarding the risk to certain groups and certain types of operations. This can happen with simple accident counts or statistical tests, for example the use of odds-ratios on accident data can give false impressions for this reason, as shown in Jarvis and Harris (2008). By treating accidents in this way, observers are in danger of drawing conclusions not conducive to safety. This was shown in chapter two (e.g. Piggott 1997, Pratt 2000). The unsupportable statements made by authors in such positions demonstrate that there are important safety reasons why observers should avoid considering accidents as populations, without factoring in all other flights.

It is recommended that the knowledge of the high accident rate to low-hours pilots is used to dispel inaccurate popular notions of the relative safety of this group, particularly among instructors. Additionally instructors require more awareness of the types of causal factors that are likely to impact upon students when they are sent solo.

As well as greater awareness, more emphasis needs to be put on the approach phase of flight in training and assessment. In particular, instructors should actively ensure that students can recognise and recover from undershoot situations. This may involve setting up such situations, but it may also be possible that the task of approach control and judgement for gliders could be taught or practiced in simulators. Research is required to establish whether good transfer of training could be achieved in this respect.

The method by which instructors and students judge the approach path requires more thought, particularly in terms of the visual references required for student learning and instructor assessment. It is recommended that reference points are used during all training flights, not just instructor demonstrations of approach control. This would help instructors to identify approach judgement issues and help students to receive feedback in order to learn the task more proficiently. However, there is also a need to establish whether the 'reference point' technique is indeed the best method of teaching this phase.

Further research

There are a number of areas that could benefit from further research on the basis of the accident findings alone. Training is one such area and the skill of approach judgement another.

The method of training approach judgement requires research to establish what is and what is not effective in terms of training. The BGA instructors' manual gives a number of 'demonstration exercises' but does not give any assistance to instructors on how to aid students in achieving effective approaches, nor on how instructors can judge when a students' judgement is adequate. This is probably because little is known about the mechanisms by which pilots perform these tasks. Although the 'reference point technique' is common in gliding literature and is the core method of teaching the approach phase (British Gliding Association 2003), it is not known whether pilots actually use such a technique in reality. It may be that as pilots gain experience they begin to acquire other ways of performing this task, particularly as reference points are not normally available. This could account for some of the accidents to early solo pilots, since they may not yet have acquired a more reliable technique. Hence research is needed to establish how pilots really perform approaches in gliders, and hence whether teaching the reference point technique is appropriate. Some of the aspects that require attention include the use of visual references, the elements that pilots pay attention to, how the information is processed, and what indicates that the glider is undershooting.

Further research is also required into instructor decision-making generally, with regard to solo flight. The decision model suggested that many such decisions failed to align with accident causes to early solo pilots, i.e. the most prominent accident cause (approach misjudgement) was unlikely to feature in the decision process. However the nature of the research meant that it was not possible to be more descriptive on this point, and so further research is required in this respect. The decision model, although only a first step, showed the possible complexity and conflict involved in the process. As stated previously, there is now a need to validate this model in the field. The model uncovered that the two mutually exclusive decision outcomes ('solo' and 'no solo') may not be completely complimentary in terms of the decision attributes that drive them. This makes it likely that heuristics and biases could play a large role in the decision process, which would mean decisions not always being based on the most critical safety factors. Since there was no way of knowing whether all aspects of the model were considered in each such decision, this point cannot be further expanded upon using the current findings. Further study is therefore required in order to uncover more about these issues. Because of the critical safety issues as well as the responsibility involved with making decisions to allow pilots to fly alone, this area requires much more research to be conducted.

References

1. AOPA (2006), *The Nall report 2006: Accident Trends and Factors for 2005*. The AOPA Air Safety Foundation 2006.

2. Andersson, B. and Nilsson, S. (1964), *Studies in the reliability and validity of the critical incident technique.* Journal of Applied Psychology. Vol 48, No 6, p.398 - 403

3. Baker, P. Lamb, W. Li, G. Dodd, R. (1996), *Crashes of Instructional Flights: Analysis of cases and remedial approaches*. Office of Aviation Medicine. Washington DC 20591. DOT/FAA/AM-96/3.

4. Bohm, J. Harris, D. (2009a), *Risk Perception and Risk-Taking Behavior of Construction Site Dumper Drivers*. International Journal of Occupational Safety and Ergonomics, In Press.

5. Bohm, J. Harris, D. (2009b), *Hazard Awareness of Construction Site Dumper Drivers*. Safety Science. Elsevier Ltd. In Press.

6. Booze, C.F. (1977), *Epidemiological investigation of occupation, age, and exposure in general aviation accidents*. Aviation, Space & Environmental Medicine 1977 Vol 48, p.1081-1091.

7. Borgen, W. A. Maglio, A. T. (2007), *Putting Action Back into Action Planning: Experiences of Career Clients*. Journal of Employment Counseling, v44 n4 Dec 2007 p.173

8. Borowsky, M. (1981), *Relationships Between Naval Aviation Safety and Pilot Flight Experience*. Aviation, Space and Environmental Medicine 1981 Vol 55, p.118-121

9. Botterill, L. Mazur, N. (2004), *Risk & risk perception. A literature review*. Rural Industries Research and Development Corporation. RIRDC Publication No 04/043.

10. Bradley, C. P. (1992), *Turning Anecdotes into Data—The Critical Incident Technique*. Family Practice 9. 1992. Oxford University Press p.98-103

11. British Gliding Association (2001), *Laws and Rules for Glider Pilots (13th ed)*. British Gliding Association, Leicester, UK

12. British Gliding Association (2003), *BGA Instructors' Manual*, (2nd ed.). The British Gliding Association Limited, Leicester, UK.

13. British Gliding Association (2004), *BGA Club Annual Statistics*. Sailplane and Gliding. 2004; Vol 55 No 2. The British Gliding Association Limited, Leicester, UK.

14. British Gliding Association (2005), *BGA Club Annual Statistics*. Sailplane and Gliding. 2005; Vol 56 No 2. The British Gliding Association Limited, Leicester, UK.

15. British Gliding Association (2006a), *BGA Club Annual Statistics*. Sailplane and Gliding. 2006; Vol 57 No 4. The British Gliding Association Limited, Leicester, UK.

16. British Gliding Association (2006b), *Glider Pilots License Notes.* http://www.gliding.co.uk/forms/gliderlicencenotes.pdf. The British Gliding Association Limited, Leicester, UK.

17. British Gliding Association (2006c), *Guidance material regarding currency*. BGA website. http://www.gliding.co.uk/forms/CurrencyGuidelinesOct06.pdf The British Gliding Association Limited, Leicester, UK.

18. British Gliding Association (2007a), *BGA Club Annual Statistics*. Sailplane and Gliding. 2007; Vol 58 No 3. The British Gliding Association Limited, Leicester, UK.

19. British Gliding Association (2007b*), UK Glider Accident Database 1997 – 2006* (Unpublished internal document), The British Gliding Association Limited, Leicester, UK. 2007

20. British Gliding Association (2008a), *Laws and Rules for glider pilots (16th Ed),* The British Gliding Association Limited, Leicester, UK

21. British Gliding Association (2008b), *Syllabus of Training: BGA solo Certificate and Bronze and cross country endorsements.* Version 4, June 2008. The British Gliding Association Limited, Leicester, UK

22. British Gliding Association (2008c), *Assistant Instructor Course Record*. The British Gliding Association Limited, Leicester, UK.

23. Bureau of Air Safety Investigation (1996), *Human Factors in Fatal Aircraft Accidents.* Department of Transport and Regional Development, Australia.

24. Butterfield, L. D. Borgan, W. A. (2005), *Outplacement Counselling from the Client's Perspective.* Career Development Quarterly, v53 n4 Jun 2005. p.306-316

25. Butterfield, L. D. Borgen, N. E. Amundson, E. A. Maglio, A. T. (2005), *Fifty Years Of The Critical Incident Technique: 1954: 2004 and Beyond*. Qualitative Research 5(4)

26. Campbell, R. Bagshaw, M. (2002), *Human Performance and Limitations in Aviation* (3rd Ed),Blackwell Science Ltd. Oxford, UK

27. Chell, E. and Pittaway, L. (1998) cited in: Gremler D (2004), *The Critical Incident Technique in Service Research*. Journal of Service Research, Volume 7, No. 1, August 2004, Sage Publications. p.65-89

28. Civil Aviation Authority (1997), CAP 667 *Review of General Aviation Fatal Accidents 1985-1994*. Civil Aviation Authority (UK)

29. Coleman, N. (2006), *Critical Incidents in Multicultural Training: An Examination of Student Experiences*. Journal of Multicultural Counseling and Development, v34 n3 Jul 2006 p.168

30. Coolican, H. (1994), *Research Methods and Statistics in Psychology*, (2nd ed.). Hodder and Stoughton.

31. Converse, P. Woolf, E. Huang, X. Oswald, F. (2008), *Response Rates for Mixed-Mode Surveys Using Mail and E-Mail/Web*. American Journal of Evaluation, v29 n1 p.99-107

32. Cottrell, D. Kilminster, S. Jolly, B. Grant, J. (2002) *What is effective supervision and how does it happen*? A critical incident study Medical Education 2002; 36: p.1042–1049

33. Dekker, S. (2006), *The Field Guide to Understanding Human Error*. Ashgate, Aldershot, UK.

34. Dollarhide, C.T. Smith, A.T. Lemberger, M.E. (2007), *Critical Incidents in the Development of Supportive Principals: Facilitating School Counselor-Principal Relationships*. Professional School Counseling, v10 n4 Apr 2007 p.360-369

35. Endsley, M. Garland, D. (2000), *Situation Awareness: Analysis and Measurement*. Lawrence Erlbaum Associates, New Jersey

36. FAA (2008), *NTSB Flight Phase Codes*. Extracted from: http://www.asias.faa.gov/aviation_studies/weather_study/fltphasecodes.html on the 1st October 2008

37. Feggetter, A. J. (1982), *A method for investigating human factor aspects of aircraft accidents and incidents*. Ergonomics, 1982, Vol. 25, No 11, p.1065-1075

38. Field, A. (2005), *Discovering Statistics Using SPSS* (2nd Ed.), SAGE Publications Ltd. London.

39. Flanagan, J. C. (1954), *The Critical Incident Technique*. Psychological Bulletin Vol. 15, No.4, July 1954

40. Fliess, J. L. (1981), *Statistical methods for rates and proportions*. New York: Wiley. Cited by: Robson, C. (2002), *Real World Research (second edition),* Blackwell Publishing, Malden.

41. Garn, A. C. Cothran, D. (2006), *The Fun Factor in Physical Education*. Journal of Teaching in Physical Education, v25 n3 Jul 2006 p.281-297

42. Gaur, D. (2005), *Human factors analysis and classification system applied to civil aircraft accidents in India*. Aviation, Space and Environmental Medicine, Vol 76, p.501–5.

43. Gerbert, K. Kemmlar, R. (1986), *The causes of causes: determinants and background variables of human factor incidents and accidents*. Ergonomics, 29, p.1439-1453.

44. Gilbert, D. T. (2002), *Inferential Correction*. Chapter 9: *Heuristics and Biases: The Psychology of Intuitive Judgement*. Gilovich, T. Griffin, D. Kahneman, D. (2002), Cambridge University Press. Cambridge. p.167-184

45. Gilovich, T. Griffin, D. (2002), *Heuristics and biases: Then and Now*. Chapter 1: *The Psychology of Intuitive Judgement*. Gilovich, T. Griffin, D. Kahneman, D. (2002), Cambridge University Press. Cambridge, p.1-18.

46. Glaser, B. G. Strauss, A. L. (1967), *The Discovery of Grounded Theory. Strategies for Qualitative Research.* Aldine Publishing Co, Chicago, IL

47. Green, R. G. Muir, H. James, M. Gradwell, D. Green, R. L. (1991), *Human Factors for Pilots*. Avebury Technical, Hants, England.

48. Gremler, D. (2004), *The Critical Incident Technique in Service Research*. Journal of Service Research, Volume 7, No. 1, August 2004, Sage Publications. p.65-89

49. Hasselquist, D.O. Baker S.P. (1999), *Homebuilt Aircraft Crashes*. Aviation, Space & Environmental Medicine. Vol 70, no 6, June 1999 p.543-547

50. Hays, W.L. (1994), *Statistics* (5th Edition), Belmont, CA: Wadsworth.

51. Heinrich, H. W. Petersen, P. E. Roos, D. B. A. (1980), *Industrial Accident Prevention, A Safety Management Approach* (5th Ed), McGraw Hill Book Company, New York.

52. Hogarth, R. (1987), *Judgement and Choice* (2nd Ed), John Wiley and Sons, New York.

53. Hollnagel, E. (1998), *Cognitive Reliability and Error Analysis Method*. Oxford: Elsevier Science Ltd.

54. Huddlestone, J. and Harris, D. (2006), *Using Grounded Theory techniques to develop models aviation of student performance*. Human Factors and Aerospace Safety 6(4), Ashgate Publishing. UK. p.371 - 382

55. Ikomi. P. Guion, R.(2000), *The Prediction of Judgment in Realistic Tasks: An Investigation of Self-Insight.* The International Journal of Aviation Psychology, 10(2), Lawrence Erlbaum Associates, Inc. p.135–153,

56. Ingstad, O. Rosness, R. Sten, T. Ulleberg, T. Rausand, M. Lydersen, S. (1990), *Helicopter Safety Study. Main Report* (STF75 A90008), Trondheim, Norway: SINTEF.

57. Jarvis, S. (2004), *Visual Sampling During the Final Turn and its effect on Flying Accuracy*. Unpublished Master's thesis. Cranfield University, Beds, UK.

58. Jarvis, S. Harris, D. (2007a), *Looking for an accident: Glider pilots' visual management and potentially dangerous final turns*. Aviation, Space and Environmental Medicine Vol 78 : p.597-600

59. Jarvis, S. Harris, D. (2007b), *Accident Rates for Novice glider pilots Vs pilots with experience.* Aviation, Space and Environmental Medicine Vol 78: p.1155-1158

60. Jarvis, S. Harris, D. (2008), *Investigation into Accident Initiation Events by Flight Phase, for Highly Inexperienced Glider Pilots*. International Journal of Applied Aviation Studies Vol 8, Number 2. FAA Academy.

61. Jenson, R. S. (1995), *Pilot Judgement and Crew Resource Management*. Ashgate, Aldershot.

62. Kahneman, D. Slovic, P. Tversky. A. (1982), *Judgement and uncertainty: Heuristics and biases*. Cambridge University Press. Cambridge.

63. Kahneman, D. Tversky. A. (1979), *Prospect Theory An analysis of Decision under risk*. Econometrica, 47:2. p.263-91.

64. Kahneman, D. Tversky. A. (2000), *Choices, Values and Frames*. Cambridge University Press. Cambride, UK.

65. Kasarskis, P. Stehwien, J. Hickox, J. Aretz, A. Wickens, C. (2001), *Comparrison of Expert and Novice Scan Behaviours During VFR Flight.* 11th International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University.

66. Keatinge, D. (2002), *Versatility and flexibility: Attributes of the Critical Incident Technique in nursing research*. Nursing and Health Sciences (2002), 4, p.33–39

67. King, N. (1998), 'Template Analysis ', cited in: Symon, G. and Cassel, C. *Qualitative methods and analysis in organisational research.* A practical guide, SAGE publications, London, p.118-134

68. Kirwan, B. Ainsworth, L. (1992), *A Guide to Task Analysis.* London: Taylor and Francis.

69. Klein, G. (1999), *Sources of Power. How people make decisions*. Massachusetts Intitute of Technology Press (MIPS), USA

70. Lenne, M. Ashby, K. (2006), *Characteristics of non-fatal general aviation crashes in Australia*. Human Factors & Aerospace Safety, 6, p.395-407

71. Li, G. (1994), *Pilot-related factors in aircraft crashes*: A review of epidemiological studies. Aviation, Space & Environmental Medicine October 1994. Vol. 65: p.944-952.

72. Li, W. Harris, D. (2005), *HFACS Analysis of ROC Air Force Aviation Accidents: Reliability Analysis and Cross-Cultural Comparison*. International Journal of Applied Aviation Studies.

73. Ludwig, D. (2005), *Use and misuse of p-values in designed and observational studies: guidance for researchers and reviewers.* Aviation, Space & Environmental Medicine Vol 76, No 7, 675-80, July 2005. p.675

74. Magill, R. A. (1989), *Motor learning, Concepts and Applications.* (3rd Ed), Wm. C. Brown Publishers. Dubuque, Iowa.

75. Neupert, K. Baughn, C. Dao, T. L. (2005), *International Management Skills for Success in Asia: A Needs-Based Determination of Skills for Foreign Managers and Local Managers.* Journal of European Industrial Training, v29 n2 p. 165-180

76. O'Hare, D. (1990), *Pilots' Perception of Risks and Hazards in General Aviation*. Aviation, Space and Environmental Medicine July 1990. Vol. 61 No. 7, p.599

77. O'Hare, D. Wiggins, M. Batt. R. Morrison, D. (1994), *Cognitive failure analysis for aircraft accident investigation*. Ergonomics; 1994, Vol 37, No.11, p.1855 - 1869

78. O'Hare, D. (2000), *The 'Wheel of Misfortune': a taxonomic approach to human factors in accident investigation and analysis in aviation and other complex systems.* Ergonomics, 2000, Vol. 43, No. 12, p.2001 - 2019

79. O'Hare, D. (2006), *Cognitive functions and performance shaping factors in aviation accidents and incidents*. The International Journal of Aviation Psychology, 16 (2), p.145-156

80. O'Hare, D. Owen, D. Wiegmann, D. (2001), *The 'where' and the 'why' of cross-country VFR crashes: database and simulation analyses*. Proceedings of the 45th Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica, CA: Human Factors and Ergonomics Society.

81. O'Hare, D. Chalmers, D. (1999), *The Incidence of Incidents: A Nationwide Study of Flight Experience and Exposure to Accidents and Incidents*. International Journal of Aviation Psychology, 9(1) p.19-32. Lawrence Earlbaum Associates, Inc.

82. Oldenburger, K. Letho, X. Feinberg, R. Letho, M. Salvendy, G. (2008), *Critical purchasing incidents in e-business*. Behaviour & Information Technology, Vol. 27, No. 1, January – February 2008, p.63 – 77. Taylor and Francis

83. Olsen, S. O. Rasmussen, J. (1989), *The reflective expert and the prenovice: Notes on skill-, rule-, and knowledge-based performance in the setting of instruction and training*. In Bainbridge, L. and Quintanilla, S. A. R. (Eds.) *Developing skills with information technology*. London: Wiley. p.9-33

84. Ostberg, O. (1980), *Risk perceptions and work behavior in forestry: implications for accident prevention policy*. Accident Analysis and Prevention 1980: 12. p.189-200

85. Padfield, G.D. Charlton, M.T. Houston, S.S. Pauser, H.J. and Hummes, D. (1988), *Observations of pilot control strategy in low level helicopter flying tasks*. Vertica Vol. 12, No.3. Pergamon Press PLC. Devon. p.203-4

86. Pagan P J, De Voogt AJ, Van Doorn RRA (2006), *Ultralight Aviation Accident Factors and Latent Failures*: A66-Case Study. Aviation, Space and Environmental Medicine Vol. 77, No. 9, p.950-952

87. Papadakis, M. (2008), *An Initial Attempt to Discover the Human Factors Underlying the Air Display Accidents*. Unpublished Master's thesis. Cranfield University, Beds, UK.

88. Patankar, M. Taylor, J. (2004), *Risk Management and Error in aviation maintenance*. Ashgate, Aldershot, UK.

89. Partington, D. (2002), *Grounded Theory*. In D.Partington (Ed) *Essential skills for management research*. London: Academic Press. p.136 - 157.

90. Piggott, D. (1991), *Gliding Safety*. A & C Black Ltd (Publishers), London.

91. Piggott, D. (1997), *Gliding. A Handbook of Soaring Flight*. A & C Black Ltd (Publishers), London.

92. Pooley, D. (2003), *Flying Training, Volume 1*. Pooley Aviation Ltd, Cranfield, UK.

93. Pratt, J. M. (1994), *The Private Pilot's Licence Course: Flying Training.* Airplan Flight Equipment Ltd, Manchester

94. Pratt, J. M. (2000), *The Private Pilot's Licence Course: Human Factors Flight Safety*. Airplan Flight Equipment Ltd, Manchester.

95. Reason, J. (1990), *Human Error*. Cambridge University Press, Cambridge, UK

96. Reichmann, H. (1993), *Cross-Country Soaring (7th Revised Edition),* Soaring Society of America Inc.

97. Renn, O. (1998), *The role of risk perception for risk management*. Reliability Engineering and System Safety, 59, p.49-62.

98. Robson, C. (2002), *Real World Research (second edition),* Blackwell Publishing, Malden.

99. Ronan, W.W. Latham GP (1974), *The Reliability and Validity of the Critical Incident Technique: A closer look*. Studies in personnel Psychology, Spring 1974, Vol. 6, No.1 p.53-64

100. Ross, L. Anderson, C.A. (1982), *Shortcomings in the attribution process: On the origins and maintenance of erroneous social assessments*. Chapter in: Kahneman, D. Slovic, P. Tversky, A. (1982), Judgement and uncertainty: Heuristics and biases. Cambridge University Press. Cambridge.

101. Rothrock, L. Yin, J. (2008), *Integrating Compensatory and Noncompensatory Decision-Making Strategies in Dynamic Task Environments*. In: Decision Modeling and Behavior in Complex and Uncertain Environments Kugler, T. Cole Smith, J. Connolly, T. Son, Y. (Ed), Springer. New York

102. Rous, B. McCormack, K. (2006), *'Critical Incident Technique: A Valuable Research Tool for Early Intervention'*. Downloaded from: http://www.ihdi.uky.edu/nectc/documents/researchSpotlight/NECTSpotlight2.pdf on 8 October 2007. Cited in Schluter, J. Seaton, P. Chaboyer, W. (2007), *Critical Incident Technique: A user's guide for nurse researchers*. Journal of Advanced Nursing 61(1), p.107–114

103. Schluter, J. Seaton, P. Chaboyer, W. (2007), *Critical Incident Technique: A user's guide for nurse researchers*. Journal of Advanced Nursing 61(1), p.107–114

104. Shappell, S. Weigmann, D. (2000), *The Human Factors Analysis and Classification System–HFACS. Final Report.* Office of aviation medicine, Washington DOT/FAA/AM-00/7.

105. Shappell, S. Weigmann, D. (2003a), *A human error analysis of general aviation controlled flight into terrain accidents occurring between 1990-1998.* Office of Aerospace Medicine, Federal Aviation Administration. Washington, DC 20591 Final Report: DOT/FAA/AM-03/4

106. Shappell, S. Weigmann, D. (2003b), *Human Error and General Aviation Accidents: A Comprehensive, Fine-Grained Analysis Using the Human Factors Analysis and Classification System.* Fiscal Year 2003 Report. Federal Aviation Administration's Human Factors Research and Engineering Division. Washington, DC 20591

107. Shappell, S. Weigmann, D. (2004), *Human Error and General Aviation Accidents: A Comprehensive, Fine-Grained Analysis Using the Human Factors Analysis and Classification System.* Fiscal Year 2004 Report. Federal Aviation Administration's Human Factors Research and Engineering Division. Washington, DC 20591

108. Simon, H. (1957), *A Behavioral Model of Rational Choice*. In: *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in Social Settings*. Wiley, New York

109. Sjöberg, L. (1987), *Risk, power and rationality: conclusions of a research project on risk generation and risk assessment in a societal perspective*. In L. Sjöberg (Ed.) Risk and Society, Allen & Unwin: London.

110. Slovic, P. (2000), *The Perception of Risk*. Earthscan Publicaions Ltd. London and Sterling, VA.

111. Slovic, P. Fischoff, B. Lichtenstein, S. (1982), *Facts versus fears: understanding perceived risk*. Chapter 33 in: Kahneman D, Slovic P, Tversky A (1982), Judgement and uncertainty: Heuristics and biases. Cambridge University Press. Cambridge. p.463-492

112. Stewart, K. (1994), *The Glider Pilot's Manual*. Airlife Publishing Ltd. England

113. Stitt-Gohdes, W.L. Lambrecht, J.J. Redman, D.H. (2000), *The Critical-Incident Technique In Job Behavior Research*. Journal of Vocational Education Research Vol 25 (1) 2000. Retreived from http://scholar.lib.vt.edu/ejournals/JVER/v25n1/stitt.html (24th September 2008)

114. Strauss, A.L and Corbin, J. (1990), *Basics of qualitative research. Grounded theory procedures and techniques.* Sage Publications, California, 1990.

115. Telfer, R. A. Ashman, A. F. (1986), *Pilot Judgment Training an Australian Validation Study.* Unpublished Research Report, Univ. of Newcastle, New South Wales

116. Telfer, R. A. (1993), *Aviation Instruction and Training*. Ashgate, England.

117. Telfer, R. A. (1989), *Pilot judgement and Decision making*. In R.S. Jenson (ed), Aviation Psychology, Aldershot England: Gower Technical

118. The Royal Society Study Group (1992), *Risk: Analysis, Perception, Management*. The Royal Society, London.

119. Thom, T. (1997), *The Air Pilot's Manual 1. Flying Training*. Airlife Publishing Limited, Shropshire, UK.

120. Tversky, A. Koehler, D. L. (1994), *Support Theory: A Nonextensional Representation of Subjective Probability*. Psychological Review, 101, American Psychological Association. p.547-567

121. van Doorn, R.R.A. De Voogt, A.J. (2007), *Glider accidents: an analysis of 143 cases, 2001-2005*. Aviation Space & Environmental Medicine. 2007; Vol 78: p.26-28.

122. van Dorn, R.R.A. Zijlstra, F. (2006), *Epidemiological study of fatal and non-fatal glider accidents in the US, 2001-2005*. Human Factors and Aerospace Safety. Ashgate, Aldershot, UK.

123. Wason, P.C. (1960), *On the failure to eliminate hypotheses in a conceptual task.* Quarterly Journal of Experimental Psychology, 12, p.129-140

124. Wells, A.T. (1992), *Flight Safety. A Primer for General Aviation Pilots*. TAB Books, A Division of McGraw Hill, US.

125. Weyman, A. K. Clarke, D. D. *Investigating the influence of organisational role on perceptions of risk in deep coal mines*. Journal of Applied Psychology 2003, 88. p.404-412

126. Wichman, H. Ball, J. (1983), *Locus of control, self-serving biases*. Aviation Space & Environmental Medicine 1983: Vol 54. p.507-10

127. Wiegmann, D. Shappell, S. (2001), *Applying the human factors analysis and classification system (HFACS) to the analysis of commercial aviation accident data*. Presented at the 11th International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University. 2001

128. Wiegmann, D. Shappell, S. (2003), *A human error approach to aviation accident analysis. A human factors analysis and classification system.* Ashgate Publishing Limited, Aldershot, UK.

129. Wiegmann, D. Faaborg, T. Boquet, A. Detwiler, C. Holcomb, K. Shappell, S. (2005), *Human Error and General Aviation Accidents: A Comprehensive, Fine-Grained Analysis Using HFACS.* Office of Aerospace Medicine, Federal Aviation Administration. Washington. Final Report No: DOT/FAA/AM-05/24

130. Woolsey, L. K. (1986), *The Critical Incident Technique. An innovative method of Research*. Canadian Journal of Counselling, Vol 20 (4)

131. Zaidman-Zait, A. (2007), *Parenting a Child with a Cochlear Implant: A Critical Incident Study*. Journal of Deaf Studies and Deaf Education, v12 n2, Spring 2007. p.221-241

Appendices

Appendix A

Guidelines for coding accidents as ‘pilot-related’ or ‘other’

Accidents must be categorised as pilot induced or non-pilot induced (other). ‘Pilot induced’ does not necessarily mean pilot culpability, responsibility or blame, but simply that the accident can be reasonably attributed to the actions taken by the pilot without an overwhelming and unavoidable influence of external or technical factors (unavoidable after boarding).

1. Pilot Related

- The causal event must have occurred AFTER the glider was boarded for flight.
- There must have been an identifiable performance shortfall in terms of the actions (or inactions) on the part of the pilot in command together with a reasonable opportunity for the pilot to act in such a way that could have avoided the accident.

2. A technical factor

- Meaning that the aircraft would have been deemed unserviceable had the failure been apparent before flight. The failure must either exist prior to the glider being boarded for flight, occur while the aircraft is within the flight envelope and operating limitations, or be induced by normal actions of the pilot in flight, while operating within the specified operating limits for the glider.
- A technical factor induced by abnormal operation of the glider (outside its operating limitations) will be deemed as pilot induced, and therefore not be counted within the technical failure category.
- A technical factor induced by the pilot on the ground (ie. rigging the glider) will still be counted as a technical factor, since for the purpose of this study ‘pilot-induced’ only refers to pilot actions after the aircraft has been boarded for flight.

Appendix A (continued)

3. An External Factor

- Any reasonably unforeseeable and reasonably unavoidable factor outside the glider, that made the flight difficult, or began the accident chain. External factors brought about by pilot actions or decisions (that were reasonably foreseeable) will be deemed to be pilot induced. For example striking a winch cable due to over-flying the winch while in operation.
- Normal weather conditions will not be counted as external factors, for example crosswind landings or cloud. Difficult flying conditions will only be counted as external factors where there were no reasonable signs or expectation of such conditions occurring.
- Lack of rising air (thermal, wave or ridge lift) will not be regarded as an external factor since such 'lift' is not reliable, is not required for safe glider operation and is not a hazardous condition for flight, unless the aircraft is put into a position such that a safe landing cannot be made.

4. General

- Accidents will only be categorised as being caused by technical and external factors where as well as the above, it is deemed that the occurrence directly led to a situation where the flight was made difficult for an average pilot. Launch failures (power failures and rope/cable breaks) will not be counted as external/mechanical factors because they are standard procedures which are trained for prior to solo and practiced during refresher and currency training. Additionally many cable/weak link breaks are induced by pilot actions.

5. Unknown

- Any accident where no causal events could be determined by the rater.

Appendix B

Guidelines for identifying and coding accident events within the accident narratives

1. Only text in the accident report descriptions and database fields should be used for the ratings.

2. The causal event must have occurred AFTER the glider was boarded for flight.

3. There must have been an identifiable performance shortfall in terms of the actions (or inactions) on the part of the pilot in command together with a reasonable opportunity for the pilot to act in such a way that could have avoided the event. This remains true even if the cause of the preceding event was deemed to be the fault of the pilot.

4. Any event identified must include a contribution, by a pilot, to the accident chain, as opposed to being a passive part in it, or a technical contribution. In other words the situation must have been worsened by the event, not just continued. For example, a low final turn as a result of a low circuit would not be included in an undershoot accident, even though it occurred between the initiating event and the crash.

5. The events must be in chronological order.

Appendices C and D

Table of all 59 causal factors with their working codes (chapter 6)

This can be used as a key for the subsequent reliability tables.

Low-level Causal factor categories		Intermediate	
J1ai	Misjudged intended separation with obstruction	Misjudged Vertical Separation	Perceptual Judgement
J1aii	Allowed wing tip to touch the ground in a turn		
J1aiii	Landing Flare too high/ too early		
J1aiv	Landing Flare too low/ too late		
J1b	Misjudged lateral separation from obstruction / object		
J2a	Flew too close in / (too high / too much energy)	Misjudged positioning in circuit	
J2b	Flew too far out/ (too low / little energy)		
J2c	Misjudged alignment of final turn exit		
J3ai	Opened airbrakes at inadvisable point in approach	Approach Misjudgement: Undershoot	
J3aii	Left airbrakes out too long		
J3aiii	Insufficient reduction of airbrakes below glide path		
J3aiv	Applied landing flap too early		
J3av	Not known		
J3bi	Used too little airbrake	Approach Misjudgement: Overshoot	
J3bii	Put airbrake away (any number of times)		
J3c	Continued until too late to safely execute alternative		
J4	Cue misinterpreted		
H1ai	Mishandled elevator causing 'longitudinal oscillation'	Mishandled in Pitch	Handling
H1aii	Overuse of up-elevator causing accelerated stall		
H1aiii	Overuse of up-elevator causing overly steep climb		
H1aiv	Overuse of up elevator causing tail strike in flare		
H1av	Overuse of down elevator		
H1ci	Not correcting/allowing for drift	Mishandled in roll	
H1cii	Mishandled sideslip		
H1d	Mishandled airbrakes		
H1e	Mishandled flaps		
H1f	Mishandled during an established aerotow		
H2a	Allowed/ failed to prevent wing going down	Ground run after landing	
H2b	Steering error (NOT misjudged lateral proximity)		
H2ci	Causing a re-light after landing by mishandling controls	Mishandled Pitch during ground run	
H2cii	Causing a premature/exaggerated take-off on launch		
H2ciii	Causing an extended ground run on take off		
H2d	Overuse of wheel brake		
H3a	Failed to maintain (or increase to) speed required	Mishandled speed	
H3b	Allowed unintentional increase of speed		
H4	Changed hands on controls		

Appendices C and D (continued)

Table of all 59 flight causal factors from chapter 6

This can be used as a key for the subsequent reliability tables

S1a	Continued into poor conditions	Continuing with a plan / strategy	Strategy
S1b	Accepted launch into unfavourable conditions		
S1c	Continued with a marginal attempt to reach airfield		
S1d	Continued operating (e.g. thermal soaring) while too low		
S1e	Flew over an area with no safe landable options		
S1f	Continued with compromised launch		
S2a	Flew out of reach of airfield during intended local flight	Flew out of reach	
S2b	Flew out of reach of chosen field while x-country/ soaring		
S3	Compromised sighting of intended flight path		
S4	Landed in an unsuitable field		
S5	Chose unsuitable initial field leading to late rejection		
S6	Chose to land on unusual/unfamiliar area of airfield		
S7	Rejected straight ahead landing from launch failure		
S8	Deliberately left wheel up		
S9	Direction of landing (into land-out field) unsuitable		
A1a	Did not notice obstruction/ditch/undulation/slope	Did not notice stimulus	Attention
A1b	Did not notice another aircraft, or launch in progress		
A1c	Did not notice change in conditions		
A2	Aircraft issue/setting overlooked		
A3	Secondary control action omitted (unintentionally)	Secondary control actions	
A4	Action initiated but not completed. Eg. u/c, canopy lock		
A5	Secondary control action slip		
A6	Failed to correctly set/lock control/straps/seat		

Appendix C part 1

Inter-rater reliability grid for reliability test (causal categories, chpr 6)

	J1a(i)	J1a(ii)	J1a(iii)	J1a(iv)	J1b	J2a	J2b	J2c	J3a(i)	J3a(ii)	J3a(iii)	J3a(iv)	J3a(v)	J3b(i)	J3b(ii)	J3c	J4	H1a(i)	H1a(ii)	H1a(iii)	H1a(iv	H1a(v)	H1c(i)	H1c(ii)	H1d	H1e	H1f	H2a	H2b	H2c(i)	H2c(ii)	H2c(iii)	H2d	H3a	H3b	H4
J1a(i)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
J1a(ii)	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1a(iii)	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
J1a(iv)	0	0	0	#	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1b	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J2a	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J2b	0	0	0	0	0	0	3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J2c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
J3a(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(iii)	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(iv)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(v)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3b(i)	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3b(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(iii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(iv)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(v)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1c(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1c(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1d	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
H1e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1f	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2a	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
H2b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
H2c(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2c(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
H2c(iii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
H2d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H3a	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	#	0	0
H3b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0
H4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
S1a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1c	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1f	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S2a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
S2b	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
S8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A1a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A1b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A1c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	S1a	S1b	S1c	S1d	S1e	S1f	S2a	S2b	S3	S4	S5	S6	S7	S8	S9	A1a	A1b	A1c	A2	A3	A4	A5	A6	C1	C2	N
J1a(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1a(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1a(iii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1a(iv)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J2a	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J2b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J2c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(ii)	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(iii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(iv)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(v)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
J3b(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3b(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J4	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(iii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(iv)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(v)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1c(i)	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1c(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1f	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2c(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2c(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2c(iii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H3a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
H3b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1a	2	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1b	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1c	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1d	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1e	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1f	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S2a	2	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
S2b	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S3	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S4	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S6	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S7	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
S8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A1a	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
A1b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A1c	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
A2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
A3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	#	0	0	0	0	0	0
A4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
A5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0
A6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0
C1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2

Appendix C part 2

Alignment of 150 events in the reliability test

Rater A	Rater B
H2a	H2a
A6	A6
S1b	S1b
J1b	J1b
J3aii	S2b
J1aiv	J1aiv
J1aiv	J1aiv
A3	A3
S2a	S1a
H2ciii	H2ciii
A1a	S4
H3a	H3a
S1b	S1b
A6	A6
A5	A5
A5	A5
J1aiv	J1aiv
H3a	J4
A5	A5
S4	S4
H1aiii	H1aiii
H3a	H3a
H3a	H3a
J3av	A2
J1aiii	J1aiii
A3	A3
S1a	S1a
J1ai	J1ai
S1d	S1d
S7	S7
J3c	J3c
S4	S4
S2a	S2a
S2a	A1c
J4	J4
J1aiv	J1aiv
A3	A3
A3	A3
J1b	J1b
A1a	A1a
J1aiii	H1aiv
S2b	S2b
A3	A3
S1b	S1b
H1d	J4
S5	S5
A3	A3
A3	A3
A2	A2
S1c	S1c

Rater A	Rater B
A3	A3
J3aiii	J3aiii
A6	A6
S3	S3
H3a	H3a
S1e	S1d
A1c	A1c
J1aiv	J1aiv
S1e	S1e
H3a	H1av
J2a	J2a
J2a	J2a
S1d	S1d
H4	H4
H1aii	H1aii
N	N
H3a	H3a
S1a	S1a
S6	S6
S2a	S1a
J2b	J3bi
A3	A3
A1c	S5
J1aiv	J1aiv
J2b	J2b
S2a	S2a
J2b	J2b
A6	A6
A3	A3
H3a	H3a
J1aiv	J1aiv
A5	A5
A5	A5
A3	A3
J1aiii	J1aiii
H2b	H2b
A3	A3
J1aiii	H3b
A1a	S4
J2b	J2b
H1ci	S1a
J1aiv	J1aiv
J4	S4
H1av	H1av
S2a	S2a
H2cii	H2cii
S4	S4
H2a	H2a
N	N
J1aiv	J1aiv

Rater A	Rater B
S3	S3
H4	H4
S4	S4
H1d	J1aiii
H1d	H1d
S1a	S3
A4	A4
S4	J4
H3a	H3a
S1c	S1c
S1c	S1c
H2a	H2a
S1d	S1d
H3a	H3a
H3a	H3a
A5	A5
S2b	J2b
A1a	A1a
S4	S4
H2a	J1aiv
H3a	A1c
H1av	H1av
H3b	H3b
J3bi	J2a
J2a	S4
S4	S4
H3a	J1ai
J2a	J2a
S1c	S1c
J2a	J2a
J2c	H1cii
A5	A5
S7	S7
H1ai	H1ai
H3b	H3b
S1c	S1c
S2a	H1ci
H2a	H2a
H1av	H1aiii
A1c	S4
J1aii	J2c
S4	S4
J1aiv	J1aiv
H3a	H3a
J1ai	H2a
H3a	H3a
J1aiv	J1aiv
S1c	J3av
H2a	H2a
S7	H2a

Appendix D part 1

Intra-rater reliability grid for reliability test

	J1a(i)	J1a(ii)	J1a(iii)	J1a(iv)	J1b	J2a	J2b	J2c	J3a(i)	J3a(ii)	J3a(iii)	J3a(iv)	J3a(v)	J3b(i)	J3b(ii)	J3c	J4	H1a(i)	H1a(ii)	H1a(iii)	H1a(iv)	H1a(v)	H1c(i)	H1c(ii)	H1d	H1e	H1f	H2a	H2b	H2c(i)	H2c(ii)	H2c(iii)	H2d	H3a	H3b	H4	S1a	S1b	S1c	S1d	S1e	S1f	S2a	S2b	S3	S4	S5	S6	S7	S8	S9	A1a	A1b	A1c	A2	A3	A4	A5	A6	C1	C2	N
J1a(i)	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1a(ii)	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1a(iii)	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1a(iv)	0	0	0	#	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J1b	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J2a	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J2b	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J2c	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(ii)	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(iii)	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(iv)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3a(v)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3b(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3b(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J3c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(iii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(iv)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1a(v)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1c(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1c(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H1f	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2c(i)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2c(ii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2c(iii)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H3a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	#	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H3b	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1e	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S1f	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S2a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S2b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
S6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0
S8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A1a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0
A1b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A1c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
A2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
A3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	#	0	0	0	0	0	0
A4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
A5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0	0	0	0
A6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0
C1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2

Appendix D part 2
Alignment of 150 events in the reliability test

Rating 1	Rating 2
H2a	H2a
A6	A6
S1b	S1b
J1b	J1b
J3aii	J3aii
J1aiv	J1aiv
J1aiv	J1aiv
A3	A3
S2a	S2a
H2ciii	H2ciii
A1a	A1a
H3a	H3a
S1b	S1b
A6	A6
A5	A5
A5	A5
J1aiv	J1aiv
H3a	H3a
A5	A5
S4	S4
H1aiii	H1aiii
H3a	H3a
H3a	H3a
J3av	J4
J1aiii	J1aiii
A3	A3
S1a	J4
J1ai	J1ai
S1d	S1d
S7	S7
J3c	J3c
S4	S4
S2a	S2a
S2a	S2a
J4	J4
J1aiv	J1aiv
A3	A3
A3	A3
J1b	J1b
A1a	A1a
J1aiii	J1aiii
S2b	S1d
A3	A3
S1b	S1b
H1d	H1d
S5	S5
A3	A3
A3	A3
A2	A2
S1c	S1c

Rating 1	Rating 2
A3	A3
J3aiii	J3aiii
A6	A6
S3	S3
H3a	H3a
S1e	S1e
A1c	A1c
J1aiv	J1aiv
S1e	S1d
H3a	H3a
J2a	J2a
J2a	J2a
S1d	S1d
H4	H4
H1aii	H1aii
N	N
H3a	H3a
S1a	S1a
S6	S6
S2a	S2a
J2b	J2b
A3	A3
A1c	S5
J1aiv	J1aiv
J2b	J2b
S2a	S2a
J2b	J2b
A6	A6
A3	A3
H3a	H3a
J1aiv	J1aiv
A5	A5
A5	A5
A3	A3
J1aiii	J1aiii
H2b	H2b
A3	A3
J1aiii	H3b
A1a	A1a
J2b	J2b
H1ci	H1ci
J1aiv	J1aiv
J4	S4
H1av	H1av
S2a	S2a
H2cii	H2cii
S4	S4
H2a	H2a
N	N
J1aiv	J1aiv

Rating 1	Rating 2
S3	S3
H4	H1ai
S4	A1a
H1d	H1d
H1d	H1d
S1a	S1a
A4	A4
S4	S4
H3a	H3a
S1c	S1c
S1c	S1c
H2a	H2a
S1d	S4
H3a	H3a
H3a	H3a
A5	A5
S2b	S2b
A1a	A1a
S4	S4
H2a	H2a
H3a	H3a
H1av	H3a
H3b	H3b
J3bi	J3bi
J2a	J2a
S4	S4
H3a	H3a
J2a	J2a
S1c	S1c
J2a	J2a
J2c	J2c
A5	A5
S7	S7
H1ai	H1ai
H3b	A1a
S1c	S1c
S2a	S2a
H2a	H2a
H1av	H1av
A1c	A1c
J1aii	J1aii
S4	S4
J1aiv	J1aiv
H3a	H3a
J1ai	J1ai
H3a	H3a
J1aiv	J1aiv
S1c	H3a
H2a	H2a
S7	S7

Appendix E

Instructor questionnaire for chapter 8

Participant information (anonymous)

Approx gliding hours:

Approx years of gliding

Instructor rating (BI, Ass Cat/ Full Cat / Examiner etc)?

Highest Gliding Badge

Age bracket (please circle) Under 20 20-29 30-39 40-49 50-59 60-69 70-79 80+

Male / Female?

This question refers to all UK gliding accidents in the last 5 years (that caused any damage or injury). Please indicate the number of P1 hours that you think were associated with the highest number of accidents in the UK (causing any damage or injury) by drawing a cross on the horizontal scale, and then writing the exact number as a figure in the box. If you think there were more than one experience level that sustained a high accident rate, please feel free to put a second cross, and a second number in the box. You can repeat this for a third time if you so wish. If you do put more than one cross, please circle the one relating to the highest number of accidents. If you feel there is a 'band' of experience associated with a high accident rate then please just indicate the lowest experience level in that band.

0 P1 Hours Increasing experience → 5000 P1 Hours

Please write the exact number (of P1 hours) in the box below (use other boxes as well if you have put more than one cross on the scale):

Now please state **very** briefly (below) why you think this experience level (or experience levels) lead to most accidents?

- Referring to all UK pilot-related gliding accidents in the last 5 years to pilots with 10 hours P1 or fewer, while solo (that caused any damage or injury):

  In what phase of flight were most accidents caused ?

- Please **rank** the following flight phases (using the numbers 6 down to 1) by which you think cause most accidents. '6' means the highest number of accident causes and '1' the lowest

Approach ☐ ☐ Circuit

Landing ☐ ☐ Launch

Pre-flight ☐ ☐ Between top of launch and start of circuit

Appendix F

Reliability matrix for coding of all participant instructors comments regarding reasons behind their estimation of the experience level associated with the highest accident rates (chapter 8).

		Original rating							
		1	2	3	4	5	6	7	
Second Judge	1	22	3	0	0	1	0	0	**26**
	2	0	10	0	4	0	0	0	**14**
	3	0	0	3	0	0	0	0	**3**
	4	0	0	0	11	1	0	0	**12**
	5	0	0	0	0	9	0	0	**9**
	6	0	0	0	0	0	6	0	**6**
	7	0	0	0	0	0	0	4	**4**
		22	**13**	**3**	**15**	**11**	**6**	**4**	**74**

- Agreement = 87.8%
- Cohen's Kappa = 0.85

Appendix G

The British Gliding Association Assistant Instructor Course Record (Course Flying Exercises). January 2008

Obtained from:
http://www.gliding.co.uk/forms/AssistantInstructorRecordJan08.pdf
4th January 2009

General
Pre flight checks
Lookout

Effects of controls
Elevator including stall
Ailerons
Rudder
Adverse yaw
ASI and speed monitoring
Trimming
Straight glide / Straight glide &scan
Turning
Slip and Skid

Type conversion
Type conversion flight
Type conversion briefing

Basic stalling
Nose Drop stall
Mush stall
Wing drop stall
Individual stall symptoms

Basic spinning / Spiral dives
Spin entry
Spin symptoms
Spin recovery
Spiral dive entry
Spiral dive symptoms
Spiral dive recovery

Further stalling
Reduced g exercise
Elevator effectiveness
High speed stall
Stall speed increases in turn

Appendix G (continued)

Further spinning

Effect of rudder near stall
Spin off thermal turn
Spin off failed launch (at height)

Circuit planning

Normal circuit
Zig Zag circuit
Low circuit (turn in early)

Aerotow

Ground roll and take off
Normal tow
Recovery from vertical displacement
Recovery from lateral displacement
Lateral instability
Recovery from divergent
Release

Winch launching

Normal launch
Low launch failure (land ahead)
Awkward' height launch failure
V.low launch failure < 50 Ft

Approach control

Airbrake elevator co-ordination
Normal approach
Undershoot and recovery
Overshoot and recovery
Progressive undershoot
Progressive overshoot

Landing

Normal landing
Balloon and recovery

Lesson planning

Pre solo training flights
Post solo check flights

Fault finding

Handling skills
judgement exercises

Appendix H
The final CIT interview Schedule

"Thank you for agreeing to take part in this study. I would like to record the interview for my notes. Within the next month I will transcribe it and then delete the recording. No one else will hear the recording, it will remain anonymous, and nothing you say will be identified. You are free to withdraw from the interview at any time and I will delete the recording on request. Once the data has been analysed, it will become completely anonymous, and so cannot be deleted after that point."

"Instructors such as yourself are constantly observing and evaluating student pilots, and we wish to tap those expertise by asking you to recall some of your own experiences related to the evaluation of student pilots for the purpose of flying solo"

"We are trying to learn in detail what standard of flying is required from instructors in order to allow a pilot to be sent solo"

"We would like you to think back to a flight or session of flights that you have had with a student pilot within the last year or so, where at some point you were engaged in a process of assessment (formal or informal) with a view to sending the student solo after the dual flight or session was complete. Importantly, if they failed, at some point before or during the session there must have been at least a chance that you would have allowed them to fly solo after the session / flight. This could be a first solo, or an early-solo check flight (ie. day checks). We would like to avoid situations of currency checks or annual checks or checks for specific purposes such as cross-country flying, aerobatics, launch or type-conversions, because we are only interested in early-solo pilots"

Eliciting the CIT responses...

1. "Think of a student that you have flown with recently, preferably within the last year, under these circumstances, but don't tell me their name."
2. "Did you allow them to fly solo that session?"
3. "What was it about that student's flying that made you say ['no' / yes] on that occasion?"

Probes

- What preceded and contributed to the 'incident'?
- What did the student do or not do that had an effect?
- What was the outcome or result?
- What made 'this action' effective or ineffective?
- What could have made the action more effective?

Activity specific probes:

- What part of the flight did it occur in?
- What was it about that student, or their flying that contributed to your decision?
- What specifically was it that they did or did not do?

"Thank you for your time, you have been very helpful. I'd like to remind you that all the information will remain anonymous and the recording will be deleted after being transcribed. Do you have any questions?"

Finally the participants were given contact details for further information.

Appendix I

Initial CIT interview schedule for piloting

Flanagan (1954) suggested forestalling any doubt in participant as to their selection by pointing out their position and the perspective that gives them on the activity. Additionally it was important to include the main CIT aim. These points accounted for the top section of the interview schedule.

The following CIT questions were arrived for the purpose of piloting.

> "*We would like you to think back to a flight or session of flights that you have had with a student pilot within the last year or so, where at some point you were engaged in a process of assessment (formal or informal) with a view to sending the student solo after the dual flight or session was complete. Importantly, prior to the flight or session there must have been at least a chance that you would have allowed them to fly solo that day. This could be a first solo or an early-solo check flight (ie. day checks). We would like to avoid situations of currency checks or annual checks or checks for specific purposes such as cross-country flying, aerobatics, launch or type-conversions, because we are only interested in early-solo pilots".*
>
> *POSITIVE CIT QUESTION (pilot): "Staying tightly within this context, we want you to talk about a time when a student pilot satisfied you that they were of solo standard. Please describe what happened"*
>
> *NEGATIVE CIT QUESTION (pilot): "Staying tightly within this context, we want you to talk about a time when a student pilot satisfied you that they were NOT of solo standard. Please describe what happened"*

While piloting the questions, a problem emerged that was found to be consistent amoung all four participants. When asked the CIT questions, all began to in talk general way about the 'solo decision' process, rather than relating actual events. De-briefings established the reasons behind this. Participants claimed that they wanted to preferred to explore the overall topic, some said they could not remember specific cases easily, and some had difficulty expressing exactly what the student did right or wrong (particularly what they did right). This further supported findings from the previous set of open interviews.

This piloting process led to the three stage interview schedule, as used.

Appendix J

Guidelines for critical incident extraction from interview transcripts

Extraction should identify any observed behaviours of the student pilot that were effective or ineffective in attaining allowance to fly solo. In other words any unit of behaviour (action, inaction, decision, communication etc) that was meaningful in its own right with regard to the decision to allow or disallow solo flight. In addition to this, any single reason given by the instructor as to why a specific student was or was not sent solo should also be included.

Issues around extracting critical incidents

1. Generality: There must be good reason to believe that the interviewee is referring to a specifically remembered case (or number of cases). Use of singular third party pronouns (he and she) is direct evidence of this, but care should be taken over the use of plural references such as ‘they’, ‘students’, people..’ etc. Such anecdotes may be deemed to come from specific events, but should be scrutinised, and referred to a second judge. Where interviewees are talking in general terms about the sort of problems that occur, or reasons that people are sent solo, these should be discounted.

2. ‘Reverse’ comments (when the interviewee mentions negative points while relating a positive anecdote, or vice-versa): Only where it is clear that events being referred to are specific and relate specifically to a pilot being allowed or not allowed to go solo, then such incidents can be included.

3. Bounding the incident: Each critical incident should contain descriptions of single behaviours or reasons that related to single skills or actions described by the instructor. For example “they got the nose down, and selected the right area to go and land in” would be two incidents because the two actions were separate phenomenon (not related in either time or skill).

4. Repeated incidents: Some interviewees repeat the same incidents a number of times. Care should be taken to recognise what is and what is not a repeat, and if doubt exists then a third judge used. If the same sort of incident occurs twice (eg. on one flight and then another, then this is not a repeat, because it would constitute two separate occurrences). Only discount a repeat if it is exactly the same event being talked about a second time. If the same specific incident is repeated then the repeats should be ignored if they contain less information than the first incident, or merged if there is more information to be gained. This prevents repeated comments being counted as two or more separate incidents.

Important: Highlighting of critical incidents: Critical incidents should be highlighted within transcripts with minimal context included around them. A different colour should be used for each incident within the same 'story' so that 'repeated' incidents can be highlighted in the same colours in order to indicate that these represent a single incident. Also any embellishments to an incident, or where an incident description is scattered within a narrative, all the parts relating to the single incident can be recognized as a single incident.

Where a part of the narrative is shared by two incidents it will be coloured grey to indicate that it is shared (or necessary context). For example where;

"The circuit was his main problem, positioning on the one hand but also his speed control was terrible"

It can be seen that the grey highlighted text is in fact required for both incidents to be understood.

Appendix K

Inter-rater check on the matching of categories from chapter 6 (working codes included) to the top level categories from the CIT ‘meaning’ analysis (indicted in cols ‘Rater A’ and ‘Rater B’)

Rater A	Rater B	Low-level Causal factor categories		Intermediate	
n11	n11	**J1ai**	Misjudged intended separation with obstruction	Misjudged Vertical Separation	Perceptual Judgement
n11	n11	**J1aii**	Allowed wing tip to touch the ground in a turn		
n11	n11	**J1aiii**	Landing Flare too high/ too early		
n11	n11	**J1aiv**	Landing Flare too low/ too late		
n11	n11	**J1b**	Misjudged lateral separation from obstruction / object		
n11	n11	**J2a**	Flew too close in / (too high / too much energy)	Misjudged positioning in circuit	
n11	n11	**J2b**	Flew too far out/ (too low / little energy)		
n11	n11	**J2c**	Misjudged alignment of final turn exit		
n11	n11	**J3ai**	Opened airbrakes at inadvisable point in approach	Approach Misjudgement: Undershoot	
n11	n11	**J3aii**	Left airbrakes out too long		
n11	n11	**J3aiii**	Insufficient reduction of airbrakes below glide path		
n11	n11	**J3aiv**	Applied landing flap too early		
n11	n11	**J3av**	Not known		
n11	n11	**J3bi**	Used too little airbrake	Approach Misjudgement: Overshoot	
n11	n11	**J3bii**	Put airbrake away (any number of times)		
n11	n11	**J3c**	Continued until too late to safely execute alternative		
n11	n11	**J4**	Cue misinterpreted		
n3	n3	**H1ai**	Mishandled elevator causing ‘longitudinal oscillation’	Mishandled in Pitch	Handling
n3	n3	**H1aii**	Overuse of up-elevator causing accelerated stall		
n3	n3	**H1aiii**	Overuse of up-elevator causing overly steep climb		
n3	n3	**H1aiv**	Overuse of up elevator causing tail strike in flare		
n3	n3	**H1av**	Overuse of down elevator		
n3	n3	**H1ci**	Not correcting/allowing for drift	Mishandled in roll	
n3	n3	**H1cii**	Mishandled sideslip		
n3	n3	**H1d**	Mishandled airbrakes		
n3	n3	**H1e**	Mishandled flaps		
n3	n3	**H1f**	Mishandled during an established aerotow		
n3	n3	**H2a**	Allowed/ failed to prevent wing going down	Ground run after landing	
n3	n3	**H2b**	Steering error (NOT misjudged lateral proximity)		
n3	n3	**H2ci**	Causing a re-light after landing by mishandling controls	Mishandled Pitch during ground run	
n3	n3	**H2cii**	Causing a premature/exaggerated take-off on launch		
n3	n3	**H2ciii**	Causing an extended ground run on take off		
n3	n3	**H2d**	Overuse of wheel brake		
n6	n6	**H3a**	Failed to maintain (or increase to) speed required	Mishandled speed	
n6	n6	**H3b**	Allowed unintentional increase of speed		
n3	n3	**H4**	Changed hands on controls		

Appendix K - continued

Inter-rater check on the matching of categories from chapter 6 (working codes included) to the top level categories from the CIT 'meaning' analysis

n5	n5	**S1a**	Continued into poor conditions	Continuing with a plan / strategy	Strategy
n5	n5	**S1b**	Accepted launch into unfavourable conditions		
n5	n5	**S1c**	Continued with a marginal attempt to reach airfield		
n5	n5	**S1d**	Continued operating (eg. thermal soaring) while too low		
n5	n5	**S1e**	Flew over an area with no safe landable options		
n5	n5	**S1f**	Continued with compromised launch		
n5	n5	**S2a**	Flew out of reach of airfield during intended local flight	Flew out of reach	
n5	n5	**S2b**	Flew out of reach of chosen field while x-country/ soaring		
n5	n5	**S3**	Compromised sighting of intended flight path		
n5	n5	**S4**	Landed in an unsuitable field		
n5	n5	**S5**	Chose unsuitable initial field leading to late rejection		
n5	n5	**S6**	Chose to land on unusual/unfamiliar area of airfield		
n5	n5	**S7**	Rejected straight ahead landing from launch failure		
n5	n5	**S8**	Deliberately left wheel up		
n5	n5	**S9**	Direction of landing (into land-out field) unsuitable		
n1	n1	**A1a**	Did not notice obstruction/ditch/undulation/slope	Did not notice stimulus	Attention
n9	n9	**A1b**	Did not notice another aircraft, or launch in progress		
n1	n9	**A1c**	Did not notice change in conditions		
n1	n1	**A2**	Aircraft issue/setting overlooked		
n1	n1	**A3**	Secondary control action omitted (unintentionally)	Secondary control actions	
n1	n1	**A4**	Action initiated but not completed. Eg. u/c, canopy lock		
n1	n1	**A5**	Secondary control action slip		
n9	n1	**A6**	Failed to correctly set/lock control/straps/seat		

Appendix L

Inter-rater reliability grid for testing the top level categories from the CIT ‘meaning’ analysis (key to category names overleaf)

Second rater

Initial Rater

	n1	n2	n3	n4	n5	n6	n7	n8	n9	n10	n11	n12	n13	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10	p11	p12	p13	p14	p15	
n1	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
n2	0	6	0	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
n3	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
n4	0	0	0	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13
n5	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
n6	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
n7	0	0	0	0	0	0	11	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12
n8	0	0	0	0	0	1	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
n9	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
n10	0	0	0	1	0	0	0	1	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
n11	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
n12	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
n13	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
p1	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
p2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	5
p3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	3
p5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	5
p6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	2
p7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	2
p9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	6
p10	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	6	0	0	1	0	0	8
p11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	5	0	0	6
p14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
p15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	6	6	4	15	2	2	13	10	6	5	2	0	0	3	6	0	2	7	1	0	2	6	6	0	0	6	0	0	110

Appendix L (continued)

Key to category names in the inter-rater reliability test on the top level categories from the CIT 'meaning' analysis

ref		Name of top-level category (prior to the check)
n1	NEGATIVE	Airmanship
n2		Airspeed control
n3		Behind the glider
n4		Circuit Planning and Judgement
n5		Consistency
n6		Gut Feeling
n7		Handling and control
n8		Indirect issues
n9		In-flight attention
n10		Not Coping
n11		Situation – Instructor conflict
n12		Soaring
n13		Strategy / Decision errors
p1	POSITIVE	Airmanship
p2		Circuit planning / Judgement
p3		Consistency
p4		Dealing Well With a Simulated Situation / Emergency
p5		'overall non-specific'
p6		Good non-technical skills
p7		Good speed control
p8		Gut Feeling
p9		Handling
p10		Indirect issues
p11		Instant Reactions
p12		Sent despite not quite right
p13		Situation-Instructor'
p14		Soaring
p15		Unusual Glider type

Appendix M

Minor changes in categorisation of CIT data made following the inter-rater testing

After the inter-rater reliability check, notes were made about any issues that appeared to arise, and the categories were scrutinised in discussion with the 2nd rater. Following these discussions, several small adjustments were made to some categories as follows:

- The ‘soaring’ category was collapsed into the ‘handling and control’ category for both positive and negative incidents.

- The category called “Sent solo despite not quite right” was collapsed into the general category. It was decided that this category did not represent a motive for sending someone solo, but more represented general comments on a person’s ability being of an acceptable standard (regardless of other issues).

- The positive category called “Good Non-Technical Skills” was dropped, and the sub-categories which belonged to it were redistributed. This category had been made up of three positive sub categories; ‘situational awareness’, ‘decisions’ and ‘judgement’. ‘Situational awareness’ was given its own category of ‘in flight attention’ (in line with the negative categorisation, since it was decided that the incidents categorised as such were expressing issues of awareness, attention issues and general comments about situational awareness). ‘Decisions’ was also given its own category (‘Decisions’). Judgement was collapsed into the ‘Circuit Planning / Judgement’ category, after looking through the incidents and discovering that all positive incidents of this type were referring to circuit planning judgement.

- Two smaller categories called ‘gut feeling’ were discussed at length. These categories represented comments from instructors expressing an intangible feeling (or discomfort) that they had about the situation. It was decided that these should be merged into the categories “Situation - instructor conflict’ and ‘situation - instructor harmony’. This resulted in a single positive and a single negative category catering for comments directly related to the flights or sessions but where general phenomenon had been expressed by reference to the instructor rather than the student. Both top level categories of ‘gut feeling’ were simply replications of the 2nd levels categories of the same name and so both remained unchanged as 2nd level categories.

- Some minor name changes were made to standardise the wording of similar positive and negative categories (e.g. “Good Speed Control” was changed to “Airspeed Control (pos)” to align with the name of the parallel negative category. All such parallel categories were given a suffix of ‘(pos)’ or ‘(neg’) in order to denote the polarity of the comment.

Appendix N

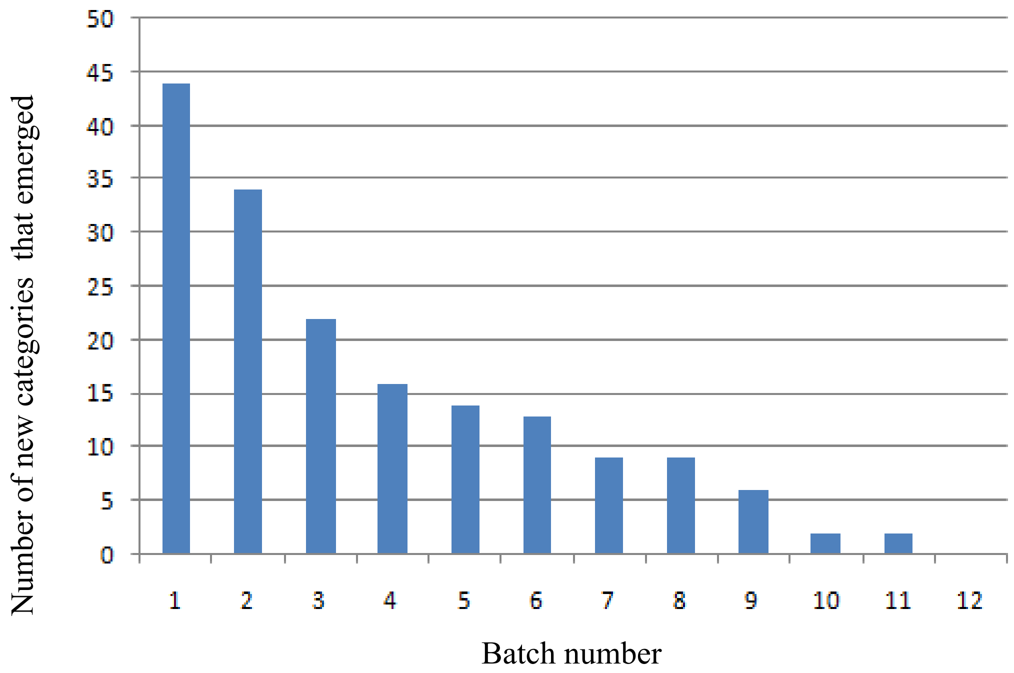
Concurrent redundancy tracking

Tracking of redundancy (of critical incidents) during data collection in order to establish coverage of the subject area and cease data collection at the appropriate time

Concurrently with the collection and categorisation of incidents, redundancy of categories was checked in order to find out when enough data had been collected to be sure that the subject area had been fully covered. After each batch (of between 50 and 63 incidents) was analysed, the number of new categories to have emerged from that batch was noted (a 'new' category was formed where an incident was found that could not be categorised with any previous incident). Twelve batches were required (each represented by a row in the table below) giving a total of 660 incidents. Each batch contained positive and negative incidents in varying amounts (for practical reasons). The total number of new categories that emerged from each batch is shown in the right-hand column of table, and can be seen to have reduced consistently until no new categories emerged from the last batch (the last 63 incidents).

Batch number	Total Number of Incidents	Number of positive incidents	Number of negative incidents	No of new categories emerging (positive)	No of new categories emerging (negative)	Total no of new categories emerging
1	58	16	42	15	29	43
2	57	29	28	16	18	34
3	53	17	36	5	17	22
4	52	16	36	8	8	16
5	59	37	22	9	5	15
6	53	17	36	5	8	13
7	58	9	49	0	9	9
8	51	5	46	1	8	9
9	55	27	28	2	4	6
10	50	11	39	0	2	2
11	51	27	24	1	1	2
12	63	17	46	0	0	0
	660	**228**	**432**	**62**	**109**	**171**

Appendix N (continued)



The above graph shows the number of new categories that emerged per batch of (approx) 50 - 60 critical incidents, and hence illustrates the trend of the redundancy. It was determined that since no new incidents emerged from the last batch, and the numbers emerging from the two preceding batches were so low, a point had been reached where no further data collection was required in order to cover the subject area. The final batch is larger than the others because even though full redundancy was found with the batch of 50 incidents, there were still some incidents remaining to be categorised at that point.

Appendix O
The 108 negative categories (from CIT analysis)

Top level categories are numbered as in Table 9.2, using N or P to represent negative or positive, then a number from 1 to 12 to represent the category. Mid level categories are numbered using a reference to the top level category to which they belong followed by their second level number (e.g. N7.1 means the first mid-level category of N7 [‘Indirect Issues’] which is ‘N7.1 - non-flying issues’). In the same way N7.1.4 is the fourth low level category of N7.1, which is N.7.1.4 - ‘Missed out a pre-flight item’. In this way, the follow tables (Appendix O and P) can be used to trace the category structure for all 170 low level categories.

Ref	Category title	n	%
N 1.1.1	drifting downwind	1	0.2
N 1.1.2	fixation on a single task	4	0.9
N 1.1.3	IP - poor situational awareness (general)	6	1.4
N 1.1.4	IP - poor situational awareness (height)	2	0.5
N 1.1.5	IP - poor situational awareness (location)	7	1.6
N 1.1.6	IP - poor situational awareness (traffic)	2	0.5
N 1.1.7	not noticing / reacting to sink	5	1.2
N 1.1.8	control slip	1	0.2
N 2.1.1	IP - behind the aircraft	8	1.9
N 2.1.2	IP - slow reactions	2	0.5
N 3.1.1	general handling poor	8	1.9
N 3.1.2	Inaccurate flying	2	0.5
N 3.1.3	IP - not flying positively	4	0.9
N 3.1.4	manner of controlling	2	0.5
N 3.1.5	Not flying smoothly	3	0.7
N 3.1.6	not keeping wings level	2	0.5
N 3.1.7	turn co-ordination poor	11	2.5
N 3.1.8	co-ordination general poor	5	1.2
N 3.2.1	Soaring (lack of ability)	1	0.2
N 3.3.1	allowing a back release	1	0.2
N 3.3.2	not allowing for drift on launch	3	0.7
N 3.3.3	Not pulling hard enough on winch	4	0.9
N 3.3.4	pulling too hard on winch	5	1.2
N 3.3.5	releasing too early	1	0.2
N 3.3.6	too steep on winch	3	0.7
N 3.3.7	aerotow handling	5	1.2
N 3.4.1	Over-controlling - nose down too far after recovery.	1	0.2
N 3.5.1	over-ruddered final turn	1	0.2
N 3.5.2	Under-banked final turn	1	0.2
N 3.5.3	Not transitioning to the approach cleanly	1	0.2
N 3.6.1	lowered nose to counter overshoot (Insufficient airbrake)	2	0.5
N 3.7.1	closing brakes on roundout	3	0.7

N 3.7.2	inappropriate use of airbrakes in circuit	2	0.5
N 3.7.3	Airbrake setting	1	0.2
N 3.7.4	airbrakes in the turn	2	0.5
N 3.8.1	over pitching in roundout	2	0.5
N 3.8.4	roundout general	5	1.2
N 3.8.5	Flying it onto the ground	3	0.7
N 3.9.1	not controlling the ground run	1	0.2
N 4.1.1	IP - gut-feel	2	0.5
N 4.2.1	IP - feel uncomfortable	2	0.5
N 4.2.2	IP - Instructor Intervention	3	0.7
N 4.2.3	IP - Nearly took over	1	0.2
N 4.2.4	IP - not doing what I would do	2	0.5
N 5.1.1	Decision making	2	0.5
N 5.1.2	IP - Perceived uncertainty	1	0.2
N 5.2.1	bad positioning after cable break	4	0.9
N 5.2.2	Land-out field choice unsuitable	1	0.2
N 5.2.3	Tried to go straight ahead after cable break	5	1.2
N 5.2.4	Turned too late after cable break	1	0.2
N 5.2.5	Turned wrong way after cable break	6	1.4
N 5.2.6	Turned, instead of going straight ahead	1	0.2
N 5.3.1	Alignment of final turn exit	2	0.5
N 5.3.2	approaching towards an obstruction / aircraft	7	1.6
N 6.1.1	Accidental stall	1	0.2
N 6.1.2	no recovery actioned	3	0.7
N 6.1.3	slow to recover speed	5	1.2
N 6.1.4	too slow final turn	1	0.2
N 6.1.5	too slow in circuit	4	0.9
N 6.1.6	too slow on the winch	1	0.2
N 6.1.7	turning with insufficient speed	2	0.5
N 6.2.1	speed not under control	22	5.1
N 6.3.1	too fast	5	1.2
N 6.3.2	too fast in circuit	7	1.6
N 7.1.1	circumstances	1	0.2
N 7.1.2	ineffective or incorrect pre-flight checks	4	0.9
N 7.1.3	knowledge poor	2	0.5
N 7.1.4	missed out a pre-flight item	3	0.7
N 7.1.5	no logbook	1	0.2
N 7.1.6	straps not done up properly	1	0.2
N 7.1.7	Based on other information	2	0.5
N 7.2.1	performance assessment	7	1.6
N 7.2.2	IP - over self critical / analytic	1	0.2
N 7.3.1	attitude issue	5	1.2
N 7.3.2	IP - over-confidence	4	0.9
N 7.3.3	IP - Panic / stress	1	0.2

N 7.3.4	IP - Physical tension	4	0.9
N 7.3.5	IP - too tired	2	0.5
N 7.3.6	IP - under-confident	5	1.2
N 7.3.7	IP - wrong frame of mind	1	0.2
N 8.1.1	General Inconsistency	9	2.1
N 8.2.1	accumulation of things	2	0.5
N 8.2.2	repeating a mistake on the next flight	2	0.5
N 9.1.1	airmanship	1	0.2
N 9.1.2	Flew too close to another glider for the instructors liking	1	0.2
N 9.1.3	lookout	13	3.0
N 10.1.1	IP - overloaded	7	1.6
N 10.1.2	IP - perceived reliance on the instructor	5	1.2
N 10.1.3	IP - wouldn't cope in emergencies	4	0.9
N 10.1.4	not coping with pressure / unusual / emergency situations	17	3.9
N 11.1.1	Base leg too close	3	0.7
N 11.1.2	circuit continuation - not aborting circuit	8	1.9
N 11.1.3	circuit planning general error(s)	8	1.9
N 11.1.4	Circuit wrong side	4	0.9
N 11.1.5	Extended base leg too far	1	0.2
N 11.1.6	extended circuit too far downwind	8	1.9
N 11.1.7	general circuit judgement poor	9	2.1
N 11.1.8	IP - Using a rigid circuit method, rather than judgement	6	1.4
N 11.1.9	judging where to joining the circuit	2	0.5
N 11.1.10	Not selecting an appropriate landing area while in circuit	3	0.7
N 11.1.11	Too close in during circuit	8	1.9
N 11.1.12	too far out in circuit	5	1.2
N 11.1.13	too high in circuit	6	1.4
N 11.1.14	too low in circuit	15	3.5
N 11.1.15	turning away from the airfield in circuit	3	0.7
N 11.1.16	unable to plan a circuit	1	0.2
N 11.2.1	Rounded out too high / early	5	1.2
N 11.2.2	Rounded out too low / late	18	4.2

Appendix O

The 62 positive categories (from CIT analysis)

Ref	Category title	n	%
P 1.1.1	situational awareness	6	2.6
P 2.1.1	instant reaction - other aircraft	1	0.4
P 2.1.2	instant reaction - sink	1	0.4
P 2.1.3	instant reaction - tug wave off	1	0.4
P 3.1.1	Accuracy / Finesse	5	2.2
P 3.1.2	co ordination good	2	0.9
P 3.1.3	general handling of the glider	4	1.8
P 3.1.4	Positive flying	1	0.4
P 3.2.1	Soaring good	1	0.4
P 3.3.1	Launch good	3	1.3
P 3.3.2	Launch safe	1	0.4
P 3.4.1	good approach control maintained despite overshoot	1	0.4
P 3.5.1	Landing good	6	2.6
P 3.5.2	Landing safe	1	0.4
P 4.1.1	instructor felt confident	5	2.2
P 4.1.2	Intangible / gut feeling	5	2.2
P 4.2.1	did what I would have done	4	1.8
P 4.2.2	Instructor feeling unnecessary	1	0.4
P 4.2.3	Instructor felt comfortable	3	1.3
P 4.2.4	Instructor not having to do or say anything	7	3.1
P 4.2.5	Instructor was not worried / uncomfortable / uneasy etc	6	2.6
P 5.1.1	Decisions good	3	1.3
P 6.1.1	Speed Control Good	8	3.5
P 7.1.1	Knowledge good	1	0.4
P 7.1.2	Pre-flight check very good	2	0.9
P 7.2.1	performance assessment	3	1.3
P 7.3.1	General impression of student	4	1.8
P 7.3.2	methodical approach	2	0.9
P 7.3.3	relaxed	2	0.9
P 7.3.4	student was confident	6	2.6
P 7.3.5	student was enthusiastic	1	0.4
P 7.4.1	solo overdue	4	1.8
P 7.4.2	Student flown a lot recently	1	0.4
P 7.4.3	Felt I had to, despite not wanting to	3	1.3
P 8.1.1	demonstrated consistency	9	4.0
P 9.1.1	Good Airmanship	1	0.4
P 9.1.2	lookout good	3	1.3
P 10.1.1	Dealt well with various simulated 'emergencies'	6	2.6
P 10.1.2	IP - will cope with emergencies	5	2.2

P 10.1.3	Sim situation - recognised the spin	1	0.4
P 10.1.4	Sim situation - recognised the stall	1	0.4
P 10.1.5	Sim situation - recovered from spin	3	1.3
P 10.1.6	Sim situation - recovered from stall	2	0.9
P 10.1.7	Sim situation - well handled launch failure	9	4.0
P 11.1.1	Circuit good	5	2.2
P 11.1.2	Circuit planning good	3	1.3
P 11.1.3	Judgement good	6	2.6
P 11.1.4	modified plan - altered circuit	3	1.3
P 11.1.5	modified plan - modified landing area	5	2.2
P 11.1.6	Recognised a problem in circuit	4	1.8
P 11.1.7	Sim situation - dealt well with simulated OOP in circuit	3	1.3
P 12.1.1	IP - competent	6	2.6
P 12.1.2	IP - safe	2	0.9
P 12.1.3	IP - student was in control	2	0.9
P 12.1.4	On top of things	2	0.9
P 12.2.1	everything was right	16	7.0
P 12.3.1	Despite problems	2	0.9
P 12.4.1	no reason not to send	5	2.2
P 12.4.2	nothing notably wrong	12	5.3
P 12.5.1	spare capacity	2	0.9
P 12.5.2	talking - pointing out what's happening	4	1.8

Appendix Q

Inter-rater reliability grid for flight phase reliability check on the CIT data.

Second rater

Initial Rater	1	2	2c	2d	2e	2f	2g	3a	3d	3g	4a	4b	4c	4d	5a	5b	6a	6b	7	8	
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2
2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2d	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2e	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
2f	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2g	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	3
3a	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	2
3d	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3g	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1
4a	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
4b	0	0	0	0	0	0	0	0	0	0	1	24	1	0	0	0	0	0	1	0	27
4c	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	3
4d	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	2
5a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
5b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6a	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11	0	0	0	11
6b	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	1	1	0	0	1	0	0	1	0	0	0	48	0	52
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	2	0	1	2	0	3	4	1	0	1	26	4	2	2	0	12	0	49	0	110